# NAVAL POSTGRADUATE SCHOOL

**MONTEREY, CALIFORNIA**

# THESIS

**EXPLORATORY SURVIVAL ANALYSIS OF DEPARTMENT OF DEFENSE BLUE-COLLAR AND WHITE-COLLAR CIVILIAN EMPLOYEE ATTRITION FACTORS**

by

Anthony C. Urech

March 2019

| | |
|---|---|
| Thesis Advisor: | Lyn R. Whitaker |
| Second Reader: | Matthew T. Geiser |

**Approved for public release. Distribution is unlimited.**

THIS PAGE INTENTIONALLY LEFT BLANK

| **REPORT DOCUMENTATION PAGE** | *Form Approved OMB No. 0704-0188* |
|---|---|



| **1. AGENCY USE ONLY** *(Leave blank)* | **2. REPORT DATE** March 2019 | **3. REPORT TYPE AND DATES COVERED** Master's thesis |
|---|---|---|

| | |
|---|---|
| **4. TITLE AND SUBTITLE** EXPLORATORY SURVIVAL ANALYSIS OF DEPARTMENT OF DEFENSE BLUE-COLLAR AND WHITE-COLLAR CIVILIAN EMPLOYEE ATTRITION FACTORS | **5. FUNDING NUMBERS** CIVATT |
| **6. AUTHOR(S)** Anthony C. Urech | |
| **7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)** Naval Postgraduate School Monterey, CA 93943-5000 | **8. PERFORMING ORGANIZATION REPORT NUMBER** |
| **9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(E**S) Office of People Analytics, Alexandria, VA | **10. SPONSORING / MONITORING AGENCY REPORT NUMBER** |

**11. SUPPLEMENTARY NOTES** The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

| **12a. DISTRIBUTION / AVAILABILITY STATEMENT** Approved for public release. Distribution is unlimited. | **12b. DISTRIBUTION CODE** A |
|---|---|

**13. ABSTRACT (maximum 200 words)**

A key enabler of military readiness includes civilian employees who work for the Department of Defense (DoD). To sustain military readiness, it is in the government's interest to understand DoD civilian workforce attrition patterns and attrition factors. The intent of this research is to better understand DoD civilian employee personnel factors that might influence attrition. To meet this intent, we use survival analysis based on calendar year 2009 new hires with covariates found in a DoD civilian's personnel record, as well as with covariates found in applicable employees' prior military active component or reserve component records. In comparison of blue-collar and white-collar employees, we see there are very similar survival trends and that retired military service members have the highest survivability. However, we do find that younger blue-collar males (29-years-old or less) have a higher survival probability than younger white-collar males, and blue-collar females have a higher survival probability than white-collar females. At the aggregate level, the probability of employee survivability increased among employees with families, higher salaries (greater than $50,000) and higher education (associate degree, bachelor's degree, or master's degree). Finally, employees who are male, who are between the ages of 35 and 54, or who work for the Navy, have an increased survivability.

| | |
|---|---|
| **14. SUBJECT TERMS** attrition, blue-collar, DoD civilian, survival analysis, white-collar | **15. NUMBER OF PAGES** 109 |
| | **16. PRICE CODE** |

| **17. SECURITY CLASSIFICATION OF REPORT** | **18. SECURITY CLASSIFICATION OF THIS PAGE** | **19. SECURITY CLASSIFICATION OF ABSTRACT** | **20. LIMITATION OF ABSTRACT** |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | UU |

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

**Approved for public release. Distribution is unlimited.**

# EXPLORATORY SURVIVAL ANALYSIS OF DEPARTMENT OF DEFENSE BLUE-COLLAR AND WHITE-COLLAR CIVILIAN EMPLOYEE ATTRITION FACTORS


Anthony C. Urech
Lieutenant Commander, United States Navy
BS, Auburn University, 2002


Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN OPERATIONS RESEARCH**

from the

**NAVAL POSTGRADUATE SCHOOL**
**March 2019**

Approved by: Lyn R. Whitaker
Advisor

Matthew T. Geiser
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

# ABSTRACT

A key enabler of military readiness includes civilian employees who work for the Department of Defense (DoD). To sustain military readiness, it is in the government's interest to understand DoD civilian workforce attrition patterns and attrition factors. The intent of this research is to better understand DoD civilian employee personnel factors that might influence attrition. To meet this intent, we use survival analysis based on calendar year 2009 new hires with covariates found in a DoD civilian's personnel record, as well as with covariates found in applicable employees' prior military active component or reserve component records. In comparison of blue-collar and white-collar employees, we see there are very similar survival trends and that retired military service members have the highest survivability. However, we do find that younger blue-collar males (29-years-old or less) have a higher survival probability than younger white-collar males, and blue-collar females have a higher survival probability than white-collar females. At the aggregate level, the probability of employee survivability increased among employees with families, higher salaries (greater than $50,000) and higher education (associate degree, bachelor's degree, or master's degree). Finally, employees who are male, who are between the ages of 35 and 54, or who work for the Navy, have an increased survivability.

THIS PAGE INTENTIONALLY LEFT BLANK

# TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

# LIST OF FIGURES

# LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

# LIST OF ACRONYMS AND ABBREVIATIONS

| | |
|---|---|
| AC | Active Component |
| AFQT | Armed Forces Qualification Test |
| CRISP-DM | Cross Industry Standard for Data Mining |
| DMDC | Defense Manpower Data Center |
| DoD | Department of Defense |
| FCS | Federal creditable service |
| FERS | Federal Employees Retirement System |
| FWS | Federal Wage System |
| GAO | Government Accountability Office |
| GS | General Schedule |
| KM | Kaplan-Meier |
| OPA | Office of People Analytics |
| OPM | Office of Personnel Management |
| PDE | Person-Event Data Environment |
| PPS | Partnership for Public Service |
| RC | Reserve Component |
| STEM | Science, technology, engineering, and math |
| TFL | TRICARE For Life |
| USCB | United States Census Bureau |

THIS PAGE INTENTIONALLY LEFT BLANK

# EXECUTIVE SUMMARY

A key enabler of military readiness includes civilian employees who work for the Department of Defense (DoD), or approximately one-third of the federal civilian workforce. To sustain military readiness, it is in the government's interest to understand DoD civilian workforce attrition patterns and attrition factors. According to Asch, Mattock, and Hosek (2014), understanding DoD civilian employee attrition is vital in light of the continuing political environment of government civilian pay freezes, furloughs, shutdowns, baby-boomer retirements, and reduced government funding for federal pensions. The intent of this research is to better understand DoD civilian employee personnel factors that might influence attrition. To meet this intent, we use survival analysis of DoD employees hired in 2009. We use covariates found in a DoD civilian's personnel record, as well as covariates found in applicable employees' prior military active component or reserve component records.

The personnel data and the analysis tools used in this research are located in a remote database server known as the Person-Event Data Environment. The majority of the data is comprised of quarterly snapshots from the Defense Manpower Data Center (DMDC). The DMDC data describe individual employees' personal attributes, such as age, years of service, and job classification. We extract from this database over 1.2 million records to construct a cohort of blue-collar and white-collar DoD civilians who began employment in 2009. After data cleaning and preparation, we analyze 62,757 observations and 20 covariates. Each observation represents an individual DoD employee and describes their work history and personal attributes that are related to their propensity to separate from DoD employment.

We use the Cross-Industry Standard Process for Data Mining (Wirth and Hipp 2000) and survival analysis. We define attrition as any employee who is no longer employed by the DoD at the end of the eight-year study period. To construct our attrition variable, we use employee transaction files that include detailed attrition reasons, such as retirement, transfer, resignation, termination, separation, or removal. We employ an additional separation category during the study—records of employees who “disappear” during the study period.

In comparison of blue-collar and white-collar employees, we see very similar trends among the survival probabilities of all the covariates analyzed between the two groups including salary, work level, military experience, annuitant status, branch of service, census bureau region, and race-ethnicity. We do, however, find that younger blue-collar males (29-years-old or less) have a higher survival probability than younger white-collar males. Blue-collar jobs are based more on experience than educational background; as a result, younger blue-collar males may be more likely to stay employed with the DoD to increase their level of expertise in a particular trade. We also find that blue-collar females have a higher survival probability than white-collar females. This may be a result of the education level differences between the two groups. Blue-collar females may be more inclined to stay with the DoD due to a lack of quality jobs in the civilian sector that only require a high school diploma.

We also find that blue-collar and white-collar retired military members have the highest survival probability of the 2009 cohort; therefore, they have high potential for being good hires for the DoD. We do, however, find that employees with no military experience tend to have greater longevity than employees who served less than 20 years in the military, based on the survival probabilities. We are not advocating for the DoD to give stronger hiring preference to military retirees, but we do find it interesting that the presence of a military background does not necessarily translate to a lower attrition rate.

Additionally, we find that 30% and 35% of blue-collar and white-collar employees, respectively, attrite during the eight-year study period and most employees who separate do so within the first two years of employment. At the aggregate level, the probability of employee survivability increases among white-collar and blue-collar employees with families, higher salaries (greater than $50,000), and higher education (associate degree, bachelor’s degree, or master’s degree). White-collar professional and administrative employees and blue-collar and white-collar employees with mid-level or senior-level jobs also have a higher survival probability. Finally, white-collar and blue-collar employees who are male, who are between the ages of 35 and 54, or who work for the Navy, have an increased survivability.

The Office of People Analytics (OPA) might use the findings of our research to better understand DoD civilian attrition factors and to implement DoD civilian employee policies to reduce attrition. The study of attrition factors may also lead to improved models and tools to better predict and forecast DoD civilian attrition.

**References**

Asch BJ, Mattock MG, Hosek J (2014) The federal civil service workforce assessing the effects on retention of pay freezes, unpaid furloughs, and other federal employee compensation changes in the Department of Defense. Report RR514, RAND Corporation, Santa Monica, CA. https://www.rand.org/pubs/research_reports/RR514.html.

Wirth R, Hipp J F (2000) CRISP-DM: Towards a standard process model for data mining. Accessed January 15, 2019, https://pdfs.semanticscholar.org/48b9/293cfd4297f855867ca278f7069abc6a9c24.pdf.

THIS PAGE INTENTIONALLY LEFT BLANK

# ACKNOWLEDGMENTS

Thank you, Megan, my lovely wife, for all your love and support during my time at the Naval Postgraduate School. Hudson and Emily, my wonderful children, thank you for being so understanding while Daddy was studying. Dr. Whitaker, thank you for taking me on as a thesis student and for all your mentorship and guidance during the thesis process. Dr. Buttrey, Dr. Brien, Dr. Carlyle, and MAJ Klingensmith, thank you for all your help and guidance. CDR Geiser, thank you for being such a great program officer and for all your support and mentorship. Scott and Michelle, thank you so much for all your support and friendship. Colette, Michael, and Jasmine from the Graduate Writing Center, thank you so much for all the grammar corrections. To my cohort, thank you everyone for all your support during the Operations Research curriculum. Finally, thank you, Sue, for all your attention to detail in getting my thesis processed.

THIS PAGE INTENTIONALLY LEFT BLANK

# I. INTRODUCTION

The 676,840 civilians employed by the Department of Defense (DoD) are vital contributors to national defense (Office of Personnel Management [OPM] 2019). The sheer number of potential civil service workforce retirements will have a tremendous impact on the federal government and, as a result, on the DoD (Shoop 2005). According to Asch, Mattock, and Hosek (2014), hiring and retaining personnel to replace these retirees is crucial to the long-term success of the DoD. Therefore, the government must understand DoD civilian workforce attrition patterns and attrition factors to ensure uncompromised military readiness. Asch et al. (2014) state that understanding DoD civilian employee attrition is vital in light of the continuing political environment of federal employee pay freezes, furloughs, shutdowns, baby-boomer retirements, and less government funding for federal pensions.

The focus of this thesis is to use DoD employee personnel data to better understand DoD civilian attrition behavior. In particular, we focus on blue-collar and white-collar employees newly appointed to permanent positions in 2009 and follow their attrition behavior over an eight-year period.

## A. INTENT OF THE RESEARCH

The intent of our research is to better understand DoD civilian employee personnel factors that influence attrition. To meet this intent, we conduct survival analysis using covariates found in each DoD civilian's personnel record, as well as covariates found in applicable employees' prior military active component (AC) or reserve component (RC) records. Examples of these covariates include age, gender, race, and military experience.

The DoD tasked the sponsor of our research, the Office of People Analytics (OPA), to analyze the large data sets consisting of DoD civilian, military AC, and military RC career paths and to monitor the effects to the DoD workforce due to political changes (OPA 2017). The Research and Analysis Center is also a partner in this research. The intent of our research is to help OPA identify covariates that influence DoD civilian attrition.

## B. BACKGROUND

The DoD civilian workforce includes employees with different types of occupational categories, appointments, and pay systems. For example, employees in our research are designated as either blue-collar or white-collar based on occupational category. We differentiate between blue-collar and white-collar employees in our study to provide a comparison for our findings and also due to the vast difference between education levels and job descriptions between the two collar types. Additionally, each employee in this study has a permanent rather than temporary appointment, which greatly affects attrition behavior. Blue-collar and white-collar federal government employees are paid using the Federal Wage System (FWS) and General Schedule (GS) pay systems, respectively. Furthermore, the cohort information used for this study consists only of employees newly appointed in 2009. We choose newly hired employees to focus the study and because we can identify with fairly high certainty new hires from their records. However, because the available DoD civilian personnel records do not contain appointment dates, we do not know how long employees hired prior to 2007 have worked for the DoD.

### 1. Occupational Category Codes

Our research involves 9,279 blue-collar and 53,478 white-collar DoD civilian employees. Collar type designators are contained in each employee's personnel file as their relevant occupational category code. The white-collar occupational category codes consist of occupational series 0001–2299 and are grouped by the following categories: "administrative," "clerical," "professional," "technical," and "other white-collar." The blue-collar category codes consist of occupational series 2501–9999 and are classified as "blue-collar." We compare blue-collar and white-collar employees to find similarities and differences in attrition behaviors between these two groups. We exclude blue-collar and white-collar employees with temporary appointments and incomplete personnel information.

### 2. Types of Appointments

The federal government hires two types of employees: permanent or temporary (OPM 2015). Permanent appointments usually include between a one-year to three-year probationary period for each employee to earn a career appointment (OPM 2015). Temporary appointments do not lead to career appointments and may only last between one and four years, although individual temporary appointments may be extended if necessary (OPM 2015).

The blue-collar and white-collar employees in this research all have permanent appointment types. We study employees with permanent appointments, rather than temporary appointments, due to the vastly higher attrition rate of temporary-appointed employees. For each appointment type, by gender and by collar type, we compute Kaplan-Meier (KM) estimates (Kaplan and Meier 1958) of the survival function (i.e., the probability that an individual is still employed by the DoD *t* years from the appointment date) and plot them in Figure 1. The KM estimates show that the overall attrition rate for calendar year 2009 newly hired, temporary-appointed DoD employees during this eight-year study is 75% compared to a 34% attrition rate for permanent-appointed employees. We also see that the likelihood of attrition in the first two years of temporary employment is 50%, and that patterns in attrition rates by gender and collar type are not the same for the two appointment types.

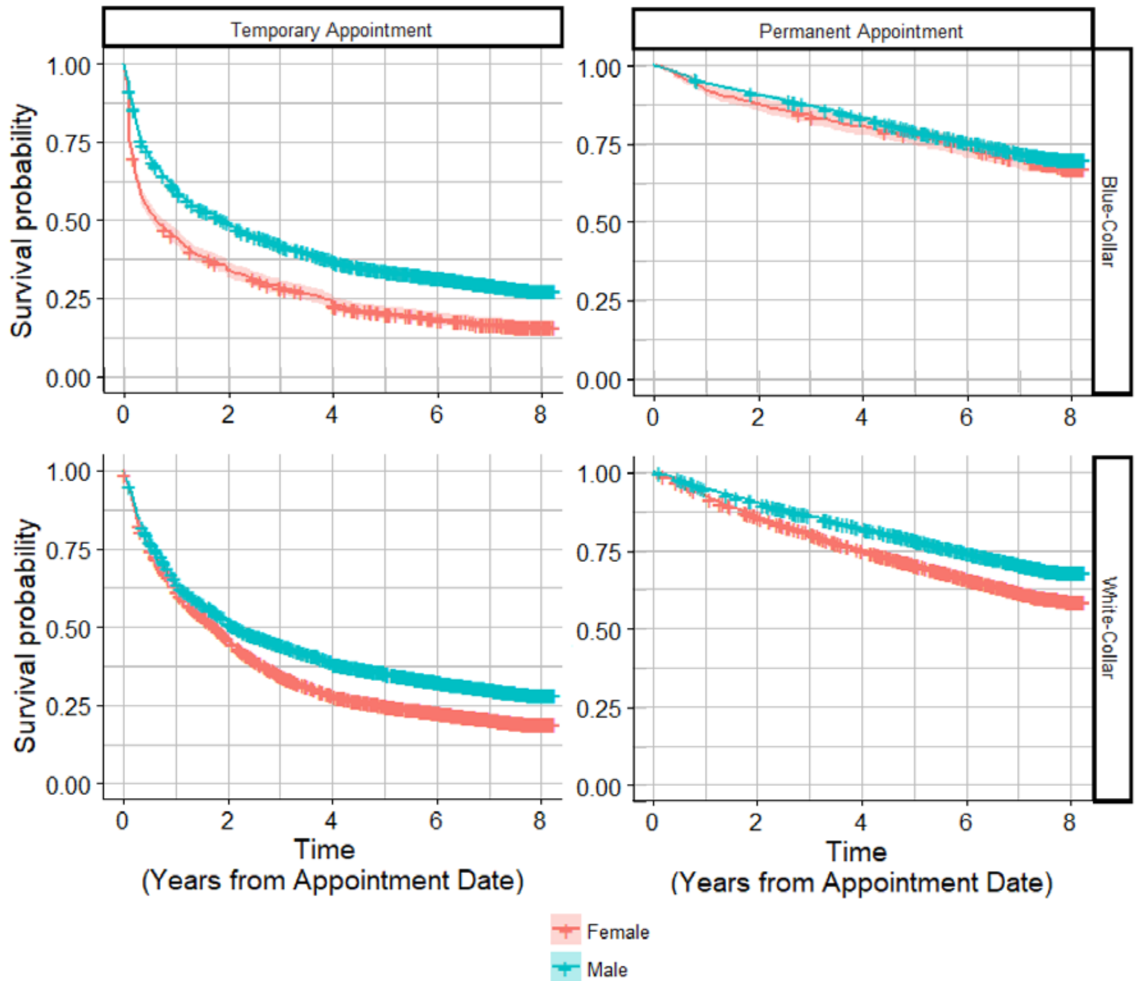



Figure 1. Kaplan-Meier (KM) Survival Function Estimates of Temporary vs. Permanent Appointment Types for DoD Employees Newly Hired in 2009 by Gender and Collar Type

### 3. Federal Wage System

The blue-collar employees in this study fall under the FWS pay system. OPM manages the FWS and is responsible for working with labor unions (OPM 2000). The FWS consists of "employees in recognized trades or crafts, or other skilled mechanical crafts, or in unskilled, semi-skilled, or skilled manual-labor occupations, and other employees including foremen and supervisors in positions having trade, craft, or laboring experience and knowledge as the paramount requirement" (OPM 2018, p. 3). According to OPM (2002), the FWS was established under law in 1972 during the presidency of Richard M.

Nixon. The law was amended in 1973 to account for appropriated-funded and nonappropriated-funded employees (OPM 2002). The system's purpose is to provide an hourly pay equivalent to blue-collar jobs in the civilian sector. There are more than 300,000 FWS employees, constituting approximately 10% of the federal civilian workforce (OPM 1999).

### 4. General Schedule Pay System

White-collar employees in this study fall under the GS pay system. "The GS classification and pay system cover the majority of civilian white-collar federal employees (about 1.5 million worldwide) in professional, technical, administrative, and clerical positions" (OPM 2015). According to OPM (2016), the GS pay system consists of paygrade GS-1 up to the paygrade of GS-15. An employee with a high school diploma may qualify only for positions at GS-4 and below, while those with either a bachelor's degree or a master's degree may qualify for positions up to GS-5 or GS-9, respectively.

## C. PREVIOUS DOD ATTRITION RESEARCH

Most available research on DoD employees focuses on AC and RC members of the military, rather than DoD civilian employees. However, we do find useful to our study federal civilian employee research conducted by the RAND Corporation (Knapp et al. 2016) and the Partnership for Public Service (PPS) (PPS 2014). We also utilize the Buttrey, Klingensmith, and Whitaker (2018) DoD civilian employee attrition study. Finally, we examine military enlistee attrition research because the cohort has a large number of prior AC and RC personnel and because of the similarities in age and education between enlisted servicemembers and blue-collar civilian employees.

### 1. RAND Corporation

RAND uses a dynamic retention model to explore how changes to DoD employee benefits and policies might influence personnel (Knapp et al. 2016). This study finds that over the span of 12 years (1998–2010) the DoD civilian workforce increased education level, diversity, and percentage of military veterans. Additionally, the research shows the increase in percentage of military veterans related directly to a rising average age of the

DoD workforce. The study also finds that military veterans have an increased desire for DoD service compared to non-veterans.

### 2. Partnership for Public Service

The PPS (2014) study of the 2013 federal workforce finds the two most common reasons for federal employee attrition are retirement or resignation. This study finds that the Army has the highest attrition rate among all government agencies and there is a "gender gap between men and women," with the 42.7% of the federal workforce that is female accounting for 43.4% of overall separations (PPS 2014). Also, due to policy changes, military veteran employment has risen within the federal workforce to more than 25% (PPS 2014). Approximately 42% of employees with "entry level" positions (GS-1 to GS-9) separate in 2013, and employees who serve less than 10 years of federal service are more likely to separate.

Although PPS (2014) describes changes in aggregate-level behaviors of federal employees, they do not study DoD employees separately; nor do they study attrition behavior as a function of the number of years employed by the DoD and other work-related factors.

### 3. DoD Civilian Attrition

Buttrey et al. (2018) examine a group of 97,654 DoD civilian employees who began employment in 2009. This study focuses on demographic factors such as age, gender, education, service component, military history, retirement eligibility, and career field (science, technology, engineering, and math [STEM] vs. non-STEM). Buttrey et al. (2018) use nonparametric survival analysis methods to study DoD civilian attrition.

Buttrey et al. (2018) find that the overall attrition rate over eight years for DoD employees who began employment in 2009 is 48%. These results show that the following DoD employees are less likely to leave their jobs: personnel with bachelor's or post-graduate degrees, STEM employees, prior AC military servicemembers (particularly those with more than 20 years of service), males, and those employed by the Navy. This study also shows that employees between the age of 21 and 49 and employees approaching retirement eligibility are more likely to remain with the DoD. Another finding of the study is the presence of a multi-modal distribution of ages, due to the high volume of employees who enter DoD civilian employment after retiring from more than 20 years in the military.

### 4. Military Enlistee Attrition

A report by the Government Accountability Office (GAO 1997) on military enlistees who entered into the military services in fiscal year 1994 shows that "about 83 percent of the 25,000 who were discharged in their first six months were assigned separation codes indicating that they (1) were medically unqualified for military service, (2) had character or behavior disorders, (3) had fraudulently or erroneously entered the military, or (4) failed to meet minimum performance criteria" (p. 4). Another report issued by the GAO (1998) used previous GAO attrition studies to show that enlistees without a high school diploma and those who earned a General Educational Development credential have higher attrition rates than those who graduated from high school. The same report also states that "those who scored in the highest category, category I, of the Armed Forces Qualification Test (AFQT) had an attrition rate of 24.7 percent, and those in category IVA had a rate of 40.7 percent" (GAO 1998, p. 3).

In his doctoral dissertation, Martin (1995) concludes that high-attrition risk recruits are males who possess at least one of the following characteristics: obesity, a record of problems with civil authorities, or less than high school education. The dissertation also finds that low-attrition risk recruits are typically either minorities (with emphasis on African-American females), females age 21 and older, males with college experience or a college degree, or those who achieve an AFQT score at or above the 65$^{th}$ percentile (with the exception of African-American males).

A complete literature review of military attrition by the U.S. Army Center for Health Promotion and Preventive Medicine (Knapik et al. 2004) identifies demographic, cognitive, and physiological factors of military enlistees who attrite and defines military attrition as "the failure of a service member to be retained in service during their contracted period" (p. 2). This report groups individuals into three-tiered education groups and shows that TIER 1 enlistees, who possess a high school diploma, have a much lower attrition rate than TIER 2 and TIER 3 enlistees, who either have a test-based diploma or did not earn a high school diploma.

Knapik et al. (2014) also examine demographic factors such as gender, age, race, ethnic group, marital status, and dependents. Their research suggests that non-African-American women, 17-or-18-year-olds, and Caucasians have greater attrition risk. Additionally, their analysis shows that marital status is not a significant factor in attrition risk; however, servicemembers with dependents have slightly lower likelihood of attrition than members without dependents.

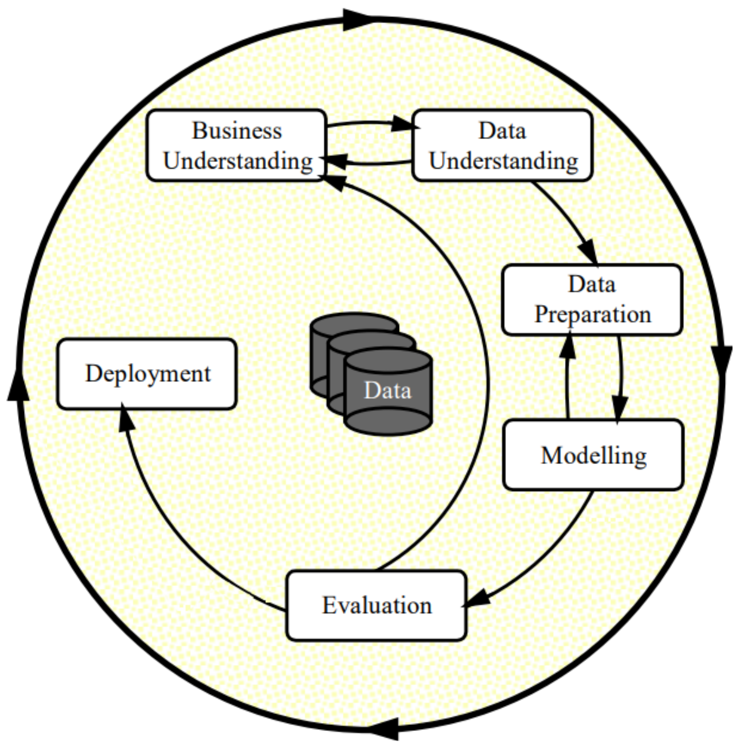
Buttrey and Clark (2017) use survival analysis to explore the propensity of enlisted sailors to “attrite, leave or reenlist at or before their first term of enlistment” (p. 1). This study uses the period from when a sailor enters the military until he or she decides to leave. By using survival analysis, they forecast what the makeup of the Navy might look like in the future based on a vast array of demographic factors.

## D. ORGANIZATION OF THESIS

The thesis has five chapters. Chapter II describes the data and methodology used to construct the 2009 cohort, a description of the datasets used in the study, and the limitations and assumptions. Chapter III presents descriptive statistics. Chapter IV contains analysis and findings, which cover the survival analyses models and their results. Chapter V concludes the thesis and provides recommendations for future work.

# II. DATA AND METHODOLOGY

In this chapter, we describe the data and the methodology that we use to conduct our research. We use the Cross-Industry Standard for Data Mining (CRISP-DM) as explained by Wirth and Hipp (2000). This chapter describes the first three phases of the CRISP-DM process, as shown in Figure 2: “business understanding,” “data understanding,” and “data preparation.” We begin by explaining some characteristics of civilian employees that might give insights into their likelihood of attrition. To more fully understand the data, we describe the datasets that include master and transaction files for civilian employees, prior AC employees, and prior-or-current RC employees. We conclude this chapter with a detailed explanation of how we construct the 2009 cohort of employees and how we build the response variable using the datasets.




Note that in this chapter we describe the first three phases of this process: “business understanding,” “data understanding,” and “data preparation.”

Figure 2. Phases of the CRISP-DM Process Model for Data Mining. Source: Wirth and Hipp (2000).

## A. BUSINESS UNDERSTANDING: THE DOD CIVILIAN EMPLOYEE

The first phase in the research methodology is to understand the DoD civilian employee. To accomplish this phase, we start with the concept map used by Buttrey et al. (2018) in their DoD civilian attrition study (see Figure 3). The concept map allows for the "organizing and representing of knowledge" due to the concept created by mapping the various relational characteristics of a DoD employee (Novak and Cañas 2008, p. 1). We modify their concept map to include the type of appointment (temporary or permanent), occupation collar type (white-collar or blue-collar), and occupational category. Using this concept map to fully understand the employee sets the groundwork for data preparation and analysis.

Many characteristics describe the DoD employee. For this research, the essential component is whether or not the employee is still working for the DoD at the end of the eight-year research period. These characteristics are listed in the concept map as "employed" or "attrited." We add one more category to employment status, that of "disappeared." These are employees whose employment records stop before the end of the study period with no indication that they have attrited. Following Buttrey et al. (2018) in our analyses, we treat these civilians as having attrited.

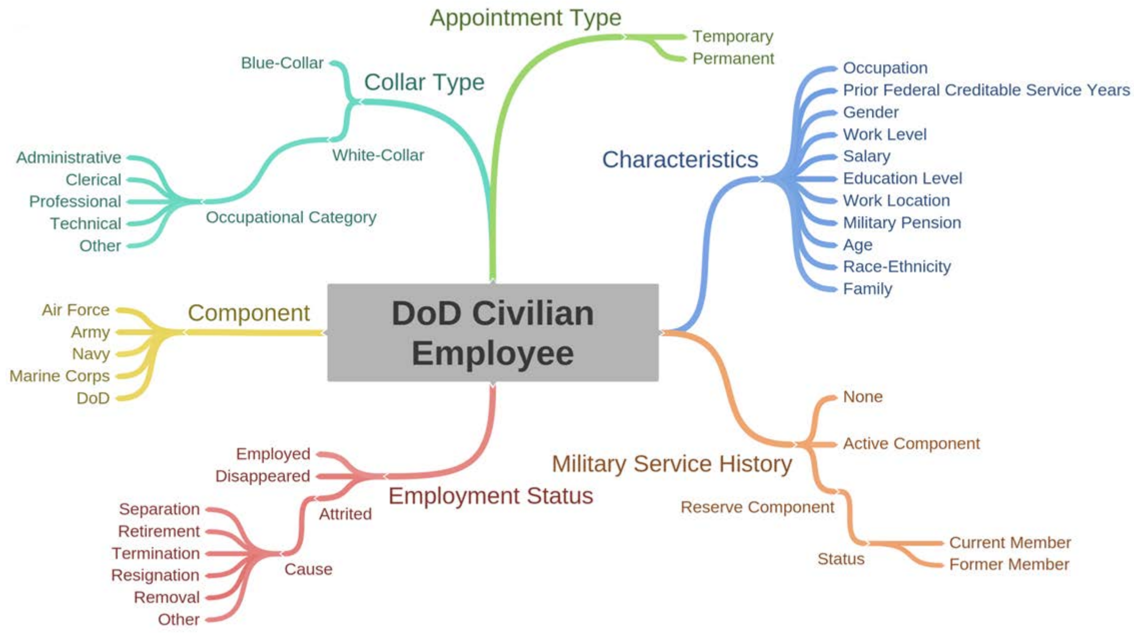



Figure 3. The DoD Civilian Employee Concept Map. Adapted from Buttrey et al. (2018).

## B. DATA UNDERSTANDING

The next phase in the research methodology is to understand the data. "The data understanding phase starts with an initial data collection and proceeds with activities to get familiar with the data, to identify data quality problems, to discover first insights into the data, or to detect interesting subsets to form hypotheses for hidden information" (Wirth and Hipp 2000, p. 5).

### 1. The Person-Event Data Environment

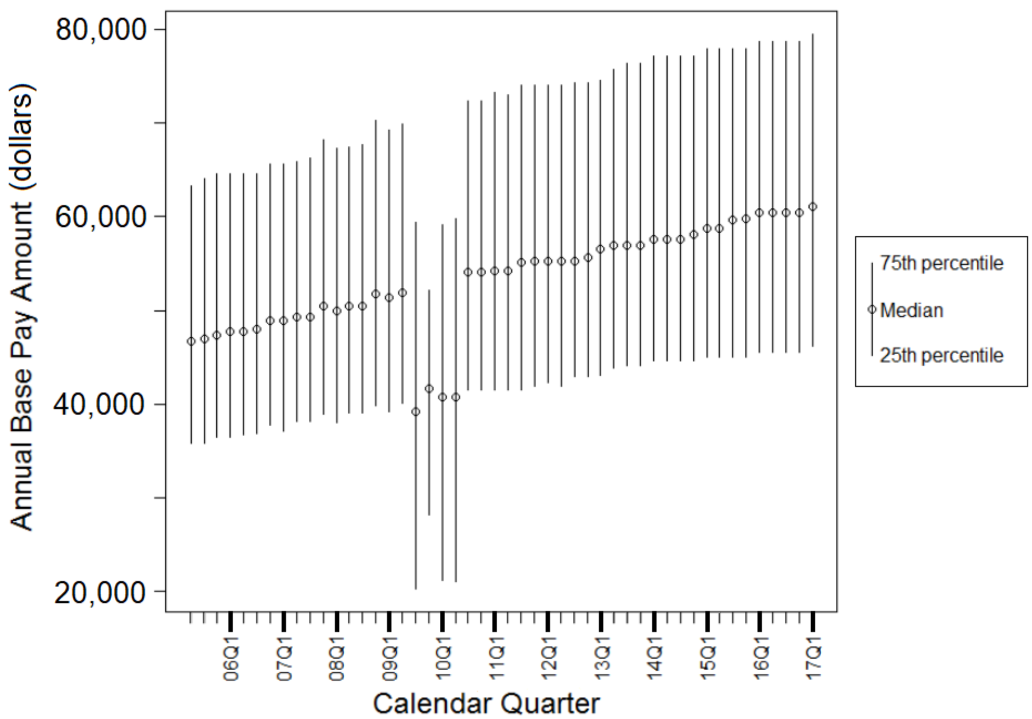
The personnel data and the analysis tools used in this research are located in a remote database server known as the Person-Event Data Environment (PDE). Knapp et al. (2018) describe the PDE and its history. In 2005, the Army partnered with the Defense Manpower Data Center (DMDC) to build the PDE.

> The purpose of the PDE was to establish an environment where research activities could share datasets for study analysis primarily in Manpower and Medical areas. The objective was to bring the analyst to the data and minimize the practice of sending data to the analyst, exposing the DoD to loss of privacy data. (DMDC 2010)

The PDE uses a network of remote Internet servers to house enormous amounts of data for various government research and projects. Access to each dataset is secure and granted only to personnel approved to work with specific datasets. In addition to providing a platform to study and analyze data using the latest statistical software tools, data in the PDE are de-identified. For example, Social Security Number identifiers are replaced with PDE generated identifiers, and all zip codes and unit identification codes are obscured or scrambled.

### 2. Data Shortcomings

Buttrey et al. (2018) discover during the construction of the original 2009 cohort that the data found in the PDE is not entirely accurate. For example, they find that 0.3% of birthdates change throughout the employee's career. They also see that the age field is only accurate 95.6% of the time. To account for this age inaccuracy, we compute an employee's age using their birthdate (month and year) given in the PDE rather than relying on pre-computed ages. They also find that the salary data for 2009 is incorrect. Figure 4 shows the calendar year quartiles of Army base pay from 2006 to 2017, with "obvious anomaly in the third and fourth quarters of 2009, and the first two quarters of 2010" (Buttrey et al. 2018, p. 30). The last shortcoming they identify is the prior service indicator, which indicates "yes" if an employee has prior military service and "no" if they do not. They discover that the prior service indicator is "yes" for all initial observations of the 2009 cohort, which is inaccurate because only 44% of the group has prior military experience.




Note that incorrect data explains the sharp change in four of 48 calendar quarters plotted.

Figure 4. The Calendar Year Quartiles of Army Base Pay. Source: Buttrey et al. (2018).

We discover further inaccuracies during our research. The first one is the incorrect race-ethnicity codes for whites and Hispanics. The race-ethnicity field is switched and uses the letters “D” and “E” to identify Hispanic and white race-ethnicity, respectively, versus “E” and “D.” Fortunately, we were able to use the Hispanic declaration code located in the PDE as a cross-reference to verify that those 2,882 employees who declare themselves Hispanic match perfectly with those coded with the letter “D.” We also use the race codes as a cross-reference to verify all other race-ethnicity codes are correct. In addition, the veteran status codes are incorrect. These codes identify employees as pre-Vietnam, post-Vietnam, or non-veterans. We find only 25% of the cohort coded as military veterans, but we know at least 44% are actual veterans based on active component (AC) and reserve component (RC) records available in the PDE. We believe this error may be due to employees who are on reserve duty, but who do not declare themselves as “veterans” because they are still serving in the military.

## C. DATA PREPARATION

The next step in the research methodology is to prepare the data. "The data preparation phase covers all activities to construct the final dataset (data that will be fed into the modeling tool(s)) from the initial raw data" (Wirth and Hipp 2000, p. 5).

### 1. 2009 DoD Civilian Employee Cohort Construction

The cohort for analysis in this research consists of 9,279 permanent-appointed blue-collar and 53,478 permanent-appointed white-collar employees newly hired in 2009 to the DoD workforce. This group is a subset of the 2009 cohort constructed by Buttrey et al. (2018), which consists of 97,876 employees of all DoD pay system and appointment types that join the DoD workforce in 2009. The 2009 cohort uses the quarterly personnel data file snapshots from all the 2009 civilian master files and also uses transaction file snapshots on the dates in which the transactions occur. No one in the 2009 group has a transaction or a master file snapshot before 2009, which indicates they are new to the DoD in the sense that they have no records between 2005 and 2009. The 2009 cohort includes newly hired employees who have worked for the DoD prior to 2005. The group does not include employees who quit while working for the DoD between 2005 and 2008 and then rejoined in 2009, because employees whose first snapshot is between 2005 through 2008 are omitted. The cohort also does not include the small number of employees with a missing birthdate. Finally, the group does not include employees with a transaction record more than 90 days prior to their first master file snapshot to ensure all the employees began employment in 2009. Figure 5 details the construction of the 2009 cohort.

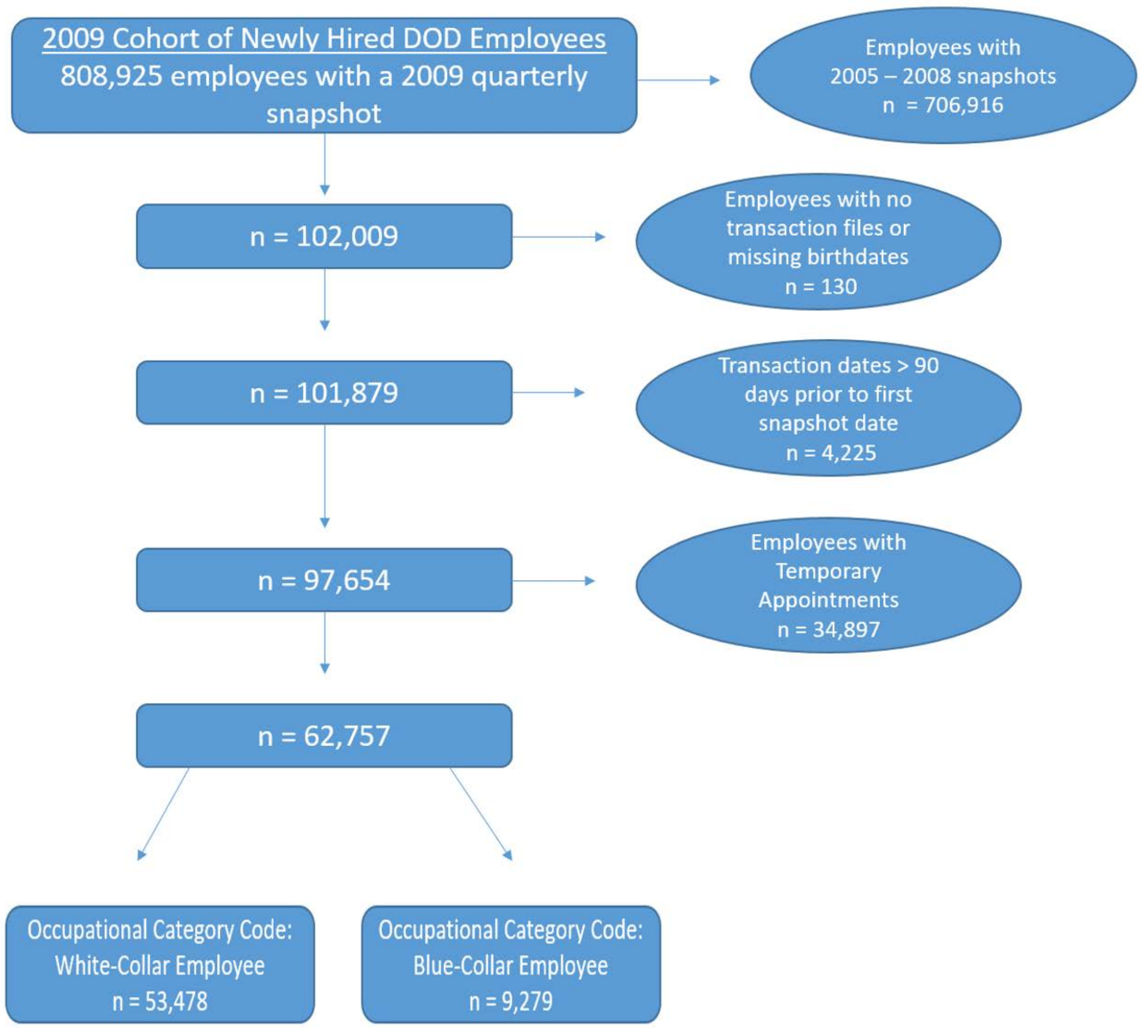



Figure 5. Flow Chart of the 2009 Cohort Construction

### 2. Datasets Used

Our research uses six types of datasets contained within the PDE. The primary datasets to build the 2009 cohort of newly hired employees are the civilian master files. These datasets, maintained by DMDC, contain demographic and detailed information found in each employee's personnel file. We merge the civilian transaction files with the civilian master files to catch all the data transactions that take place within each employee's record. These transactions include changes in the employee's career, such as salary changes, changes of appointment, and, most importantly, separation. We flag those employees with separation transactions to determine which employees attrite during the eight years of our study. We also classify employees as "disappeared" who do not have a separation transaction on file, but whose master file quarterly snapshots end during the research period, thereby indicating they are no longer employed.

Next, we add the records from the AC master file. The AC master file contains dependent quantity, dependent type, and marital status information, which is not present in the civilian master or transaction files. The data also includes the AFQT scores and enlisted career status codes. We then add the AC transaction files to this subset of prior AC employees to gain insights about their discharge from AC military service.

Lastly, we add the records from the RC master and transaction files. We merge these files to gain information about the employees who had served or are currently serving in the RC upon entering employment. These files also contain dependent quantity and marital status information, but no dependent-type information. Also included are AFQT scores and a prior AC service indicator. We also use the RC transaction files to gain insights about their discharge from RC military service.

### 3. Covariates Used

After combining records from all six dataset types, the covariates we choose to analyze include 20 categorical covariates and numeric variables. The numeric variables of age and prior federal creditable service (FCS) years are transformed into categorical variables. The snapshot dates for all covariates take place at or before the employee's first civilian transaction or master file snapshot date, which avoids bias in our survival analysis model by not using data from the future after the employee begins employment. The covariates in Table 1 are divided into two types, "time-constant" and "time-varying." Time-constant means the covariate values remain unchanged for more than 85% of the cohort throughout the eight-year study period compared to time-varying covariates, which change over time for more than 85% of the cohort. We choose these covariates based on DoD military and civilian attrition research, and we also select covariates that we perceive might be possible attrition factors, such as health and life insurance coverage plans. Table 1 provides the covariate name, type, source, and number of factor levels.

Table 1. Covariate, Type, Data Source, and Factor Levels

| Covariate | Type | Data Source | Factor Levels |
|---|---|---|---|
| Age Group | Time-Constant | Master (Civilian) | 11 |
| AFQT Category | Time-Constant | Master (AC, RC) | 6 |
| Annuitant Status Code | Time-Constant | Master (Civilian) | 3 |
| Branch of Service | Time-Constant | Master (Civilian) | 5 |
| Bureau of the Census Division Code | Time-Constant | Master (Civilian) | 9 |
| Education Level | Time-Varying | Master (Civilian) | 6 |
| Federal Group Life Insurance Program | Time-Constant | Master (Civilian) | 3 |
| Federal Creditable Service Years | Time-Constant | Master (Civilian) | 6 |
| Gender | Time-Constant | Master (Civilian) | 2 |
| Healthcare Plan | Time-Varying | Master (Civilian) | 4 |
| Inter-Service Separation Code | Time-Constant | Transaction (AC, RC) | 3 |
| Military Experience | Time-Constant | Master (Civilian, AC, RC) | 3 |
| Military Rank Grouping | Time-Constant | Master (AC, RC) | 6 |
| Occupational Category Code | Time-Constant | Master (Civilian) | 6 |
| Occupational Family | Time-Constant | Master (Civilian) | 21 |
| Occupational Group | Time-Constant | Master (Civilian) | 28 |
| Race-Ethnicity Code | Time-Constant | Master (Civilian) | 5 |
| Salary | Time-Varying | Master (Civilian) | 4 |
| Work Level | Time-Varying | Master(Civilian) | 3 |
| Years of AC Service | Time-Constant | Master (AC) | 5 |

### 4. Attrition Variable and Categories

We use survival analysis to study the distribution of time until DoD employee separation. OPM (2014) states that "separations are actions that end employment with an agency" (p. 3). Separation transactions are found in the nature action codes located in the employee transaction files and include detailed separation reasons such as retirement, transfer, resignation, termination, separation, or removal. An additional separation category created during the study is that of employees who "disappear" during the eight-year study period. The employees who "disappear" are those for whom master and transaction file snapshots end prior to 2017, but for whom there is no indication of separation in the last transaction record available (Buttrey et al. 2018).

The variable, "event," is constructed to be "true" if the employee separates or disappears during the study, and "false" if he or she is still employed at the end of the study period. Figure 6 details the distribution of the attrition categories of the blue-collar and white-collar cohorts by gender. Similar to the PPS (2014) study, we also find the majority separation reason is "resignation" for both genders and collar types. We also see approximately 25% of separated employees have the separation reason of "disappear" for each gender and collar type, so we limit ourselves in the ability to accurately account for attrition reasons based on these civilian employee snapshots that "disappear" during the study (Buttrey et al. 2018).

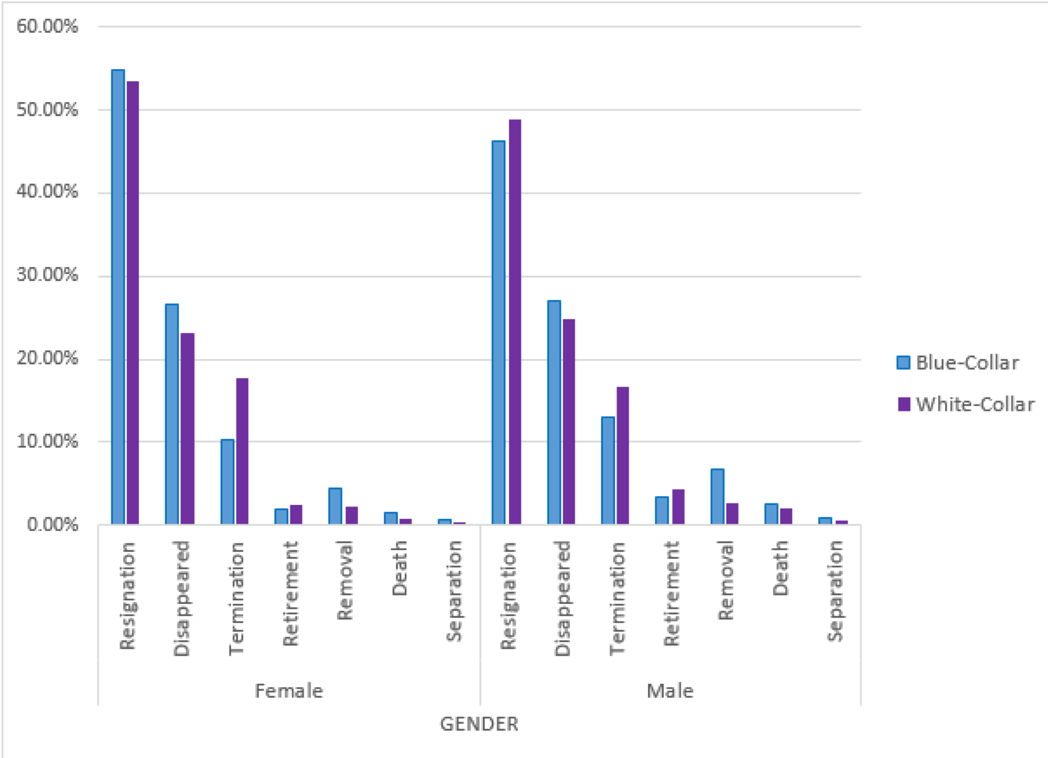



Figure 6. Distribution of Attrition Categories by Gender and Collar Type

## D. LIMITATIONS AND ASSUMPTIONS

The crucial limitation of the study is the quality of the data (Buttrey et al. 2018). Transaction records are available only from 2007. Taking a conservative approach, to allow the collection of transaction records to become established if needed, we choose to begin with the 2009 employee records, which affords us only eight years of data to study employee behavior. The eight-year period is not long enough to analyze the complete career path of civilian employees, because the average career lasts 13 years (OPM 2013). However, the eight-year range of data provides insights into early DoD civilian employee attrition behavior. We also note that the two-year economic recession, which ended in June of 2009, may have some impact on the attrition behavior of the 2009 employees (Goodman and Mance 2011).

We are also missing the valuable data fields of marital status, the number of dependent children, and work location zip codes, which limits our ability to thoroughly analyze the employees. Additionally, 4,500 employees, or 7%, have snapshot records that "disappear" before the end of the study period. We treat these employees as having separated from the DoD.

Further, we define a new hire to be those employees whose "first" master file snapshot is in 2009 and whose earliest transaction file date is in the quarter of their first snapshot date, even though we have no master file snapshot records prior to 2005 and no transaction file records prior to 2007.

# III. DESCRIPTIVE STATISTICS

This chapter offers descriptive statistics of the 9,279 blue-collar and 53,478 white-collar employees. We begin with the almost time-constant covariates of gender, age, race-ethnicity, Bureau of the Census division code associated with place of employment, branch of service, number of creditable years of federal service, life insurance coverage, and occupational descriptions, including occupational category, group, and family. We then turn our attention to “time-varying” covariates, including education level, salary, work level, and health insurance. Finally, we present the time-constant covariates that relate to prior military AC and RC employees, including military experience, annuitant status, military paygrade, AC years of service, AFQT scores, and military separation reasons.

## A. TIME-CONSTANT COVARIATES

In this section, we describe multiple nearly time-constant covariates. We treat these covariates as “time-constant” in our study because they change over time for less than 15% of the cohort or, in the case of age, they change at a constant rate throughout the study. Table 2 shows the breakdown of employee records by the first snapshot date for use in the study.

Table 2. Number of Records by First File Date for 2009 Cohort

| 2009–03–31 | 2009–06–30 | 2009–09–30 | 2009–12–31 |
|---|---|---|---|
| 13,825 | 15,370 | 20,267 | 13,295 |

### 1. Gender

The majority of blue-collar employees in the DoD are men. The Equal Opportunity Employment Commission (EEOC) report (EEOC 2009) shows that only 10.64% of the federal blue-collar workforce is female. Figure 7 shows that 9.5% of newly hired blue-collar employees are women, which is strikingly lower than the 38% of the white-collar cohort. These percentages do not uphold the EEOC report (2009) that shows the 2009 federal workforce consists of 44.1% women.

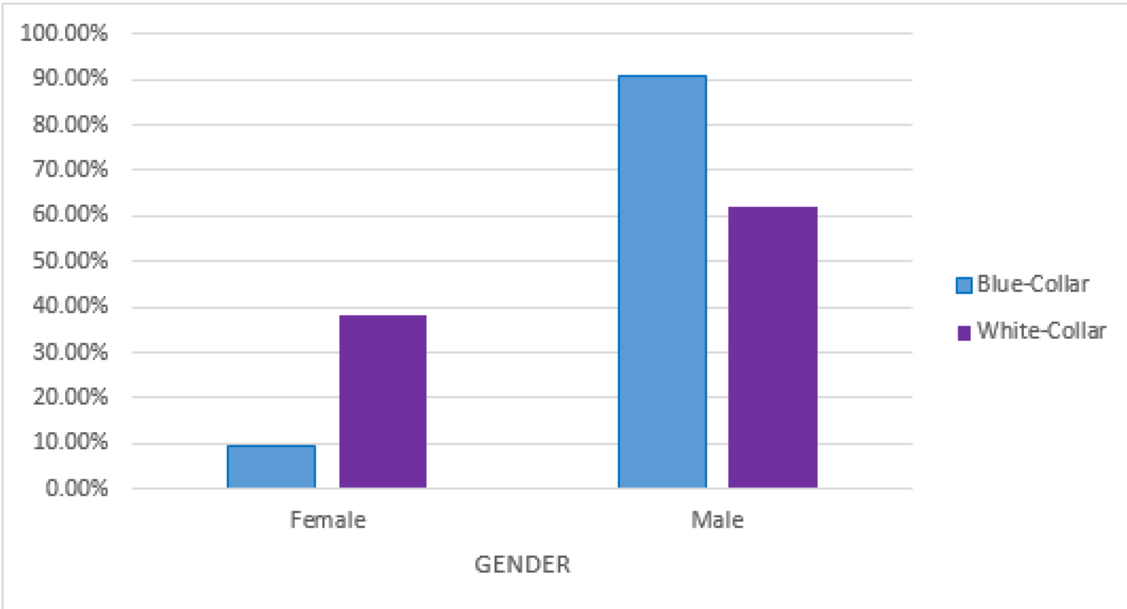



Figure 7. Gender Percentage by Collar Type

### 2. Age

The employees range from 16 to 76 years of age. The median age of female blue-collar and white-collar employees is 31 and 36, respectively, while the median age of male blue-collar and white-collar employees is 36 and 40, respectively. In Figure 8, we present the age distribution by each of the five services. We see that the Air Force hires many young blue-collar employees age 30 or under, which is needed for the aging DoD workforce. We also notice the bi-modal age distribution among both collar types between the Army, the Navy, and the Air Force, which is likely a result of these services hiring more retired AC employees who are in their late-thirties to mid-forties. This retired AC bi-modal age distribution is readily apparent in Figure 9, where we illustrate the age distribution between prior AC employees, RC employees, and employees without military experience.

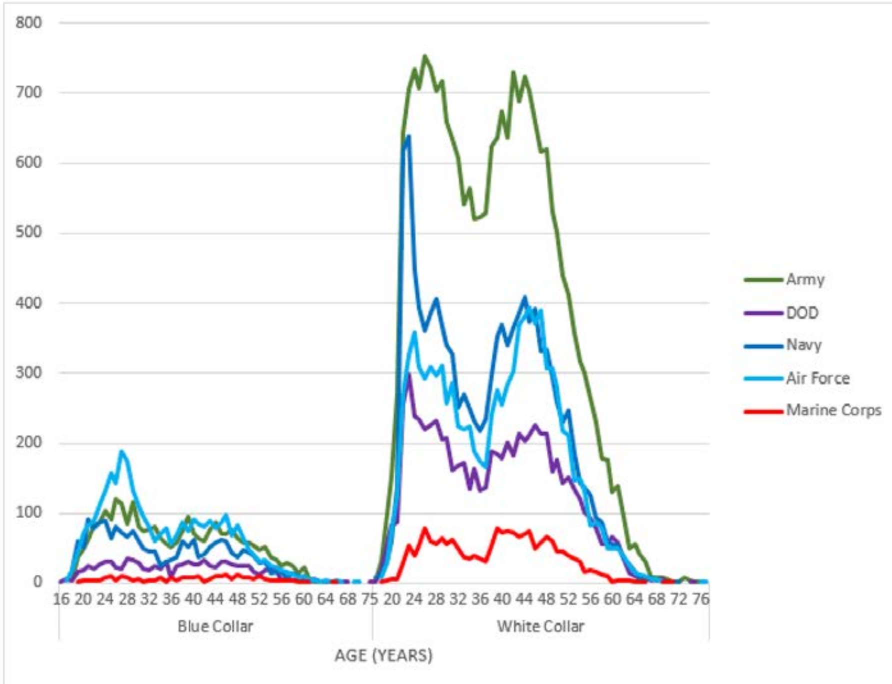



Figure 8. Distribution of Ages by Branch of Service

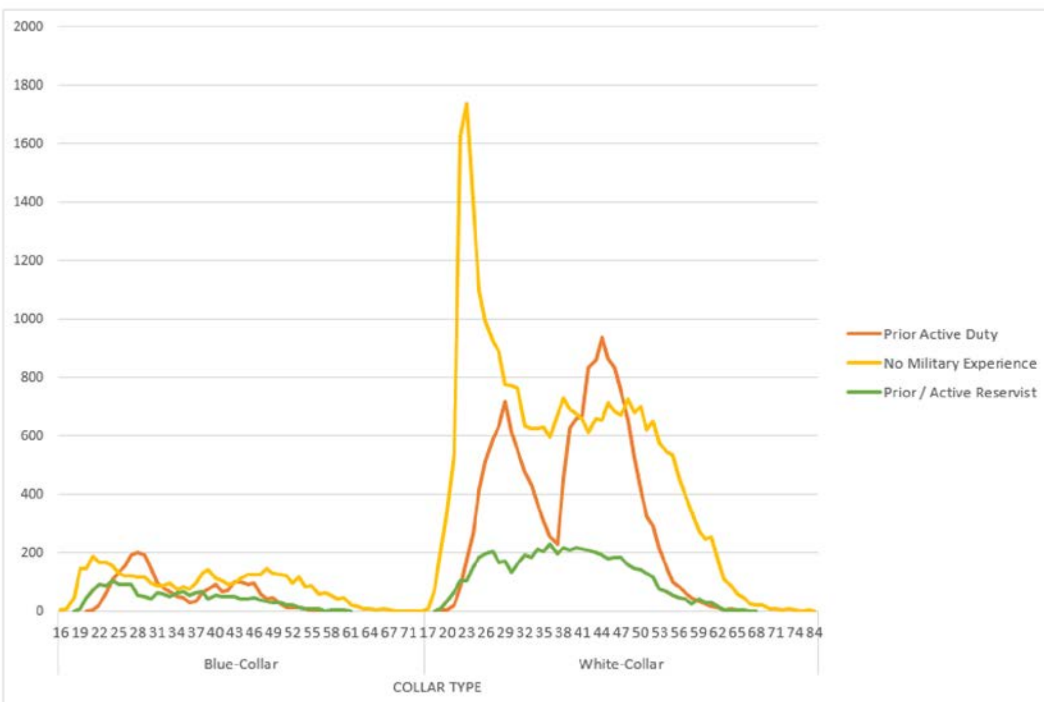



Figure 9. Age (Years) at Appointment by Prior Military Experience

We divide employees into eleven age group categories. The blue-collar portion of the workforce has a higher percentage of employees under the age of 29 than the white-collar portion. Younger blue-collar employees do not come as a surprise, because a majority of the blue-collar workforce has a high school diploma only and therefore may enter the workforce at an earlier age than typical college graduates. Figure 10 presents the distribution by age group across gender and collar types.

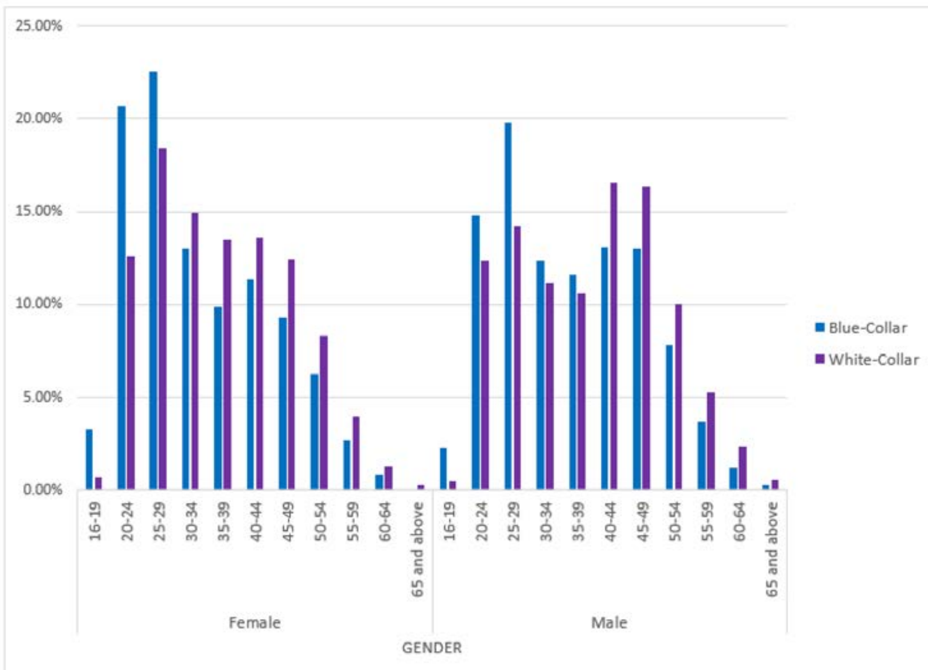



Figure 10. Age (Years) Group Distribution by Gender and Collar Type

### 3. Race-Ethnicity

The employees are divided into five race-ethnicity categories including Native American or Native Alaskan, Hispanic, Asian or Pacific Islander, Black, and White. We find the majority of the group is White among both genders and collar types, with American Indian or Native Alaskan as the smallest minority. We see that the female workforce is more ethnically diverse than the male workforce and has very similar proportions to the race-ethnicity of the EEOC report (2009), which shows that the 2009 federal government comprises 65.6% Whites, 18.0% Blacks, 7.9% Hispanics or Latinos, 6.1% Asian or Pacific

Islanders, 1.7% American Indian or Native Alaskans, and 0.7% mixed races. Figure 11 displays the diversity of the cohort by race-ethnicity and gender.

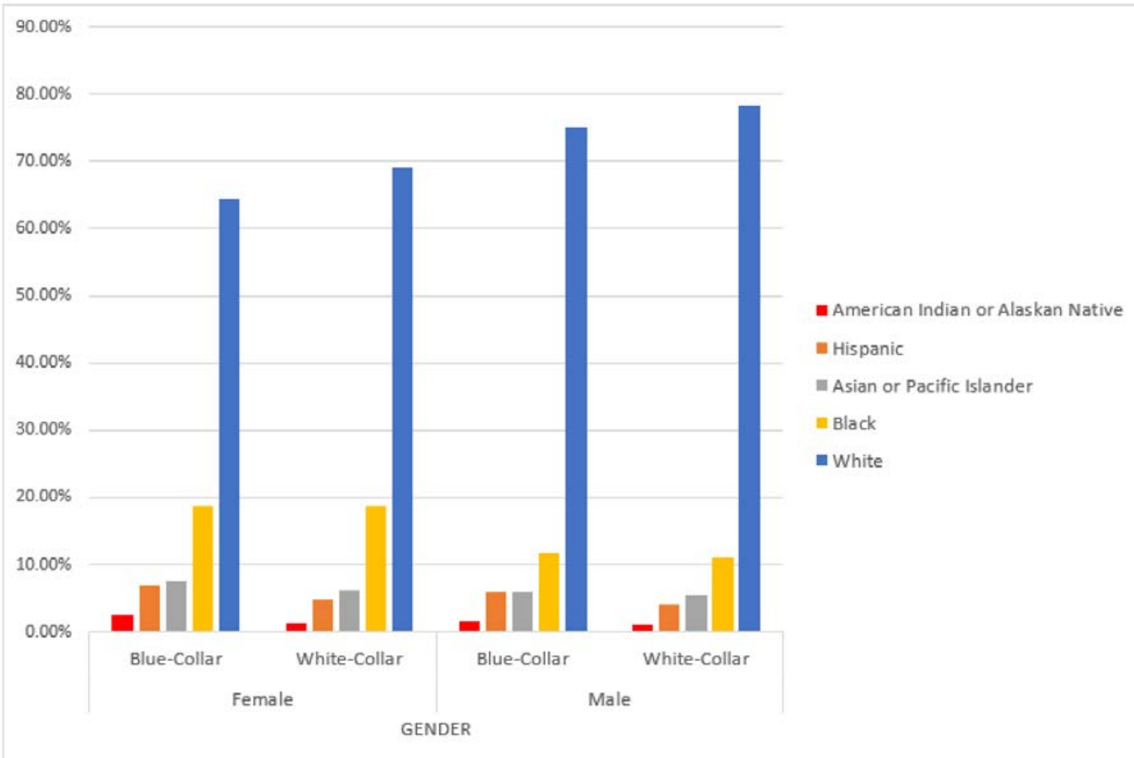



Figure 11. Race-Ethnicity by Gender and Collar Type

### 4. Bureau of the Census Division

The employees work across the country in nine locations that correlate to the United States Census Bureau (USCB) division codes. We also add the location of "overseas" for employees who do not fall within the USCB divisions. A detailed map of these divisions is in Appendix A. Figure 12 shows the distribution of employees across the USCB divisions, and we see that the South Atlantic division has the most employees across all genders and collar types.

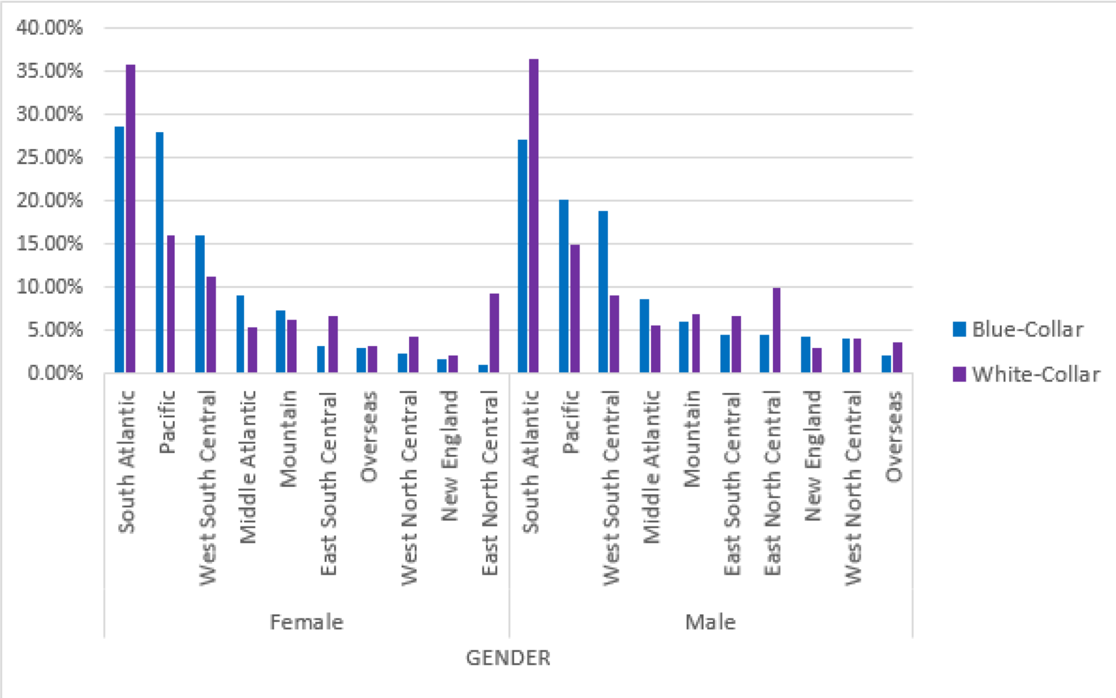



Figure 12. Distribution of USCB Divisions by Gender and Collar Type

### 5. Branch of Service

The employees work for either the DoD, the Army, the Navy, the Air Force, or the Marine Corps. Fewer than 5% of employees switch services during the study, so we focus on the service in which they begin their employment. We see that the Army is the largest employer of the white-collar employees and the Air Force hires the majority of blue-collar employees regardless of gender, as shown in Figure 13. We also see that the Marine Corps hires the fewest employees, which makes sense as it is the smallest of the group.

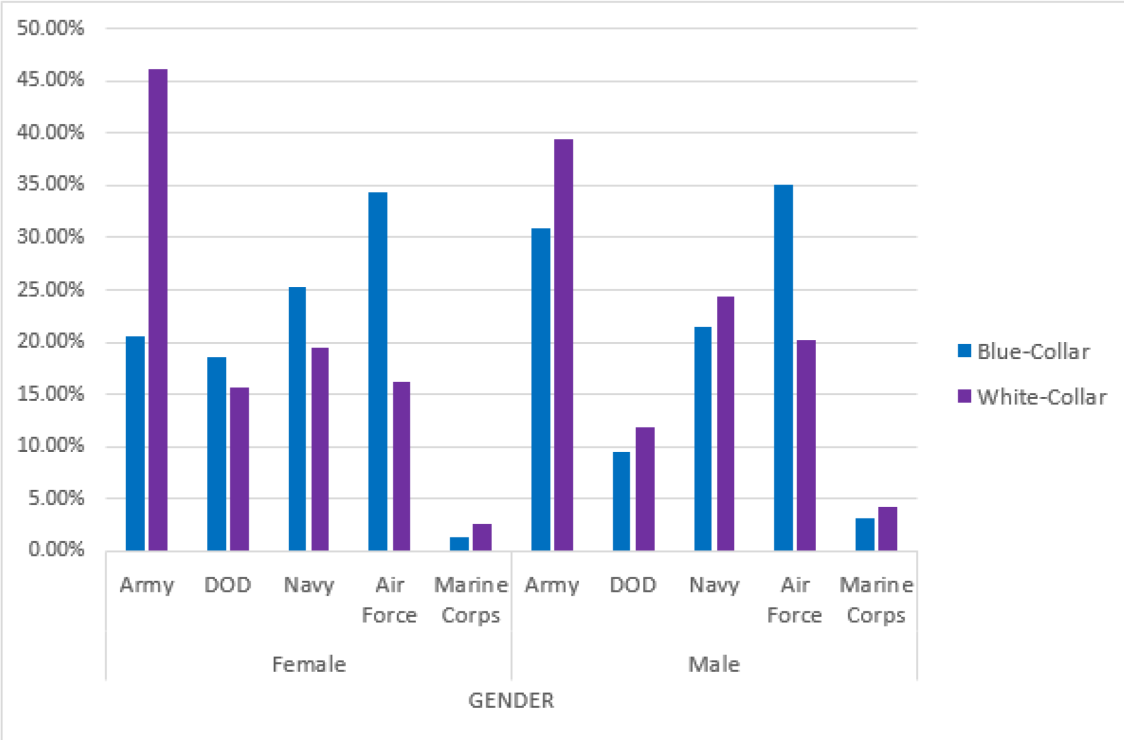



Figure 13. Distribution of Branch of Service by Gender and Collar Type

#### 6. Prior Federal Creditable Service Years

The Federal Employees Retirement System (FERS) is available to most blue-collar and white-collar employees. To receive any retirement benefit, employees must complete at least five years of federal creditable service (FCS) and they may retire if they complete 20 or 30 years of service depending on their desired retirement age (OPM 2010). We group FCS years into five categories: none, 1–4, 5–10, 11–15, 16–19, and 20 and above. Figure 14 shows the distribution of FCS years by age and collar type. We see that the majority of new hires have no FCS years. Our intuition is that those employees with 15 or more completed FCS years are less of an attrition risk because they are much closer to earning a pension than those employees with no or fewer completed FCS years.

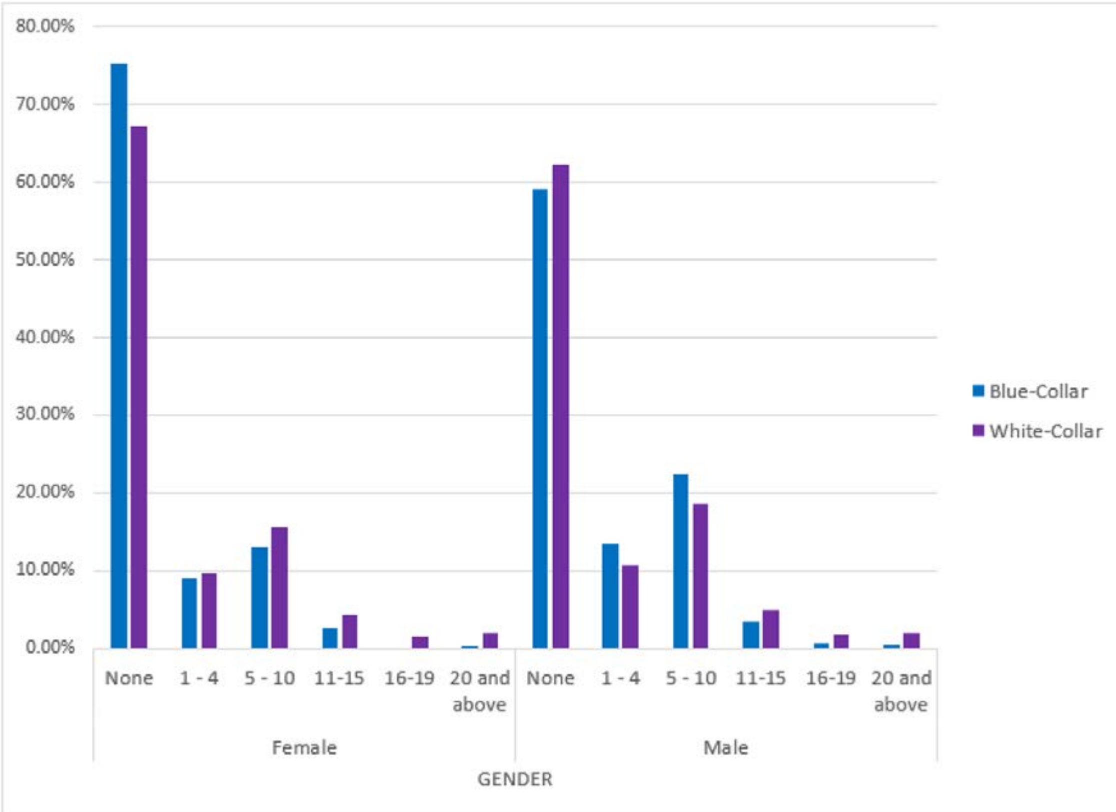



Figure 14. Distribution of FCS Years by Gender and Collar Type

### 7. Occupational Categories

The employees are divided into five occupational categories that represent the overall class of an employee's position. Occupational categories of "administrative," "clerical," "professional," "technical," and "other" apply to white-collar employees. The blue-collar occupational category is "blue-collar" and refers to trade, craft, and labor occupations (OPM 2009). Professional work requires a bachelor's or advanced degree in a specific field (OPM 2009). Administrative work requires "knowledge of one or more fields of administration or management" (OPM 2009, p. 9). Technical work "involves extensive practical knowledge, gained through experience and/or specific training less than that represented by college graduation" (OPM 2009, p. 10). Clerical positions "involve structured work in support of office, business, or fiscal operations" (OPM 2009, p. 10). The "other" occupational category encompasses positions that do not fit into the different categories and includes occupations such as police and firemen. (OPM 2009). Figure 15 shows the distribution of occupational categories by gender, and we see that the majority for white-collar females and males is “professional” and “administrative,” respectively.

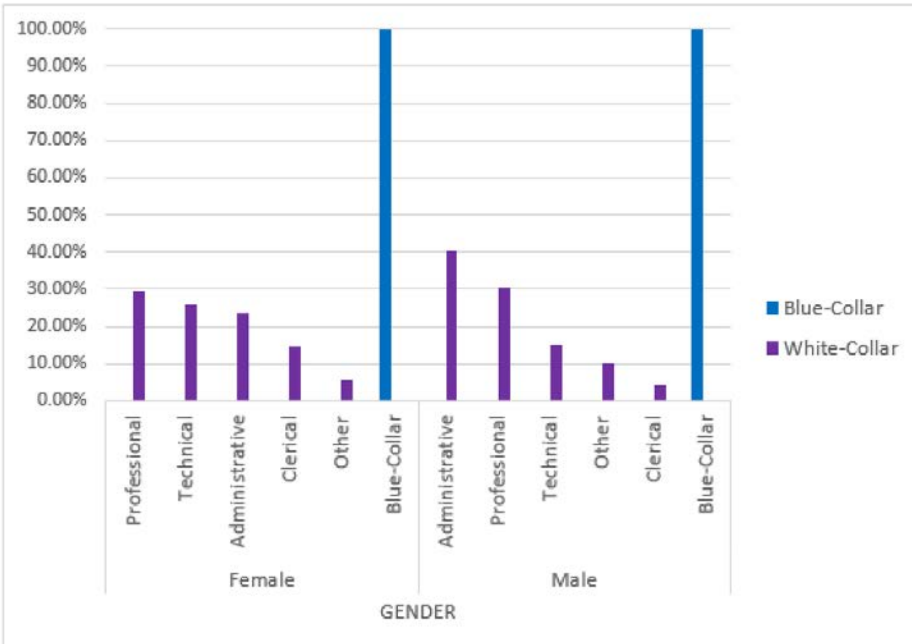



Figure 15. Distribution of Occupational Categories by Gender and Collar Type

### 8. Occupational Groups

White-collar jobs in the federal government are broken down into 23 occupational groups (OPM 2018). Within these groups are occupational series and codes that list the employee’s specific position. Due to the numerous variations of employee positions and the limited number of employees in each particular area, we focus our study on the occupational groups rather than series. Figure 16 shows the cohort’s distribution of white-collar occupational groups by gender. We notice the majority occupational group for both genders is the “General Administrative, Clerical, and Office Services.”

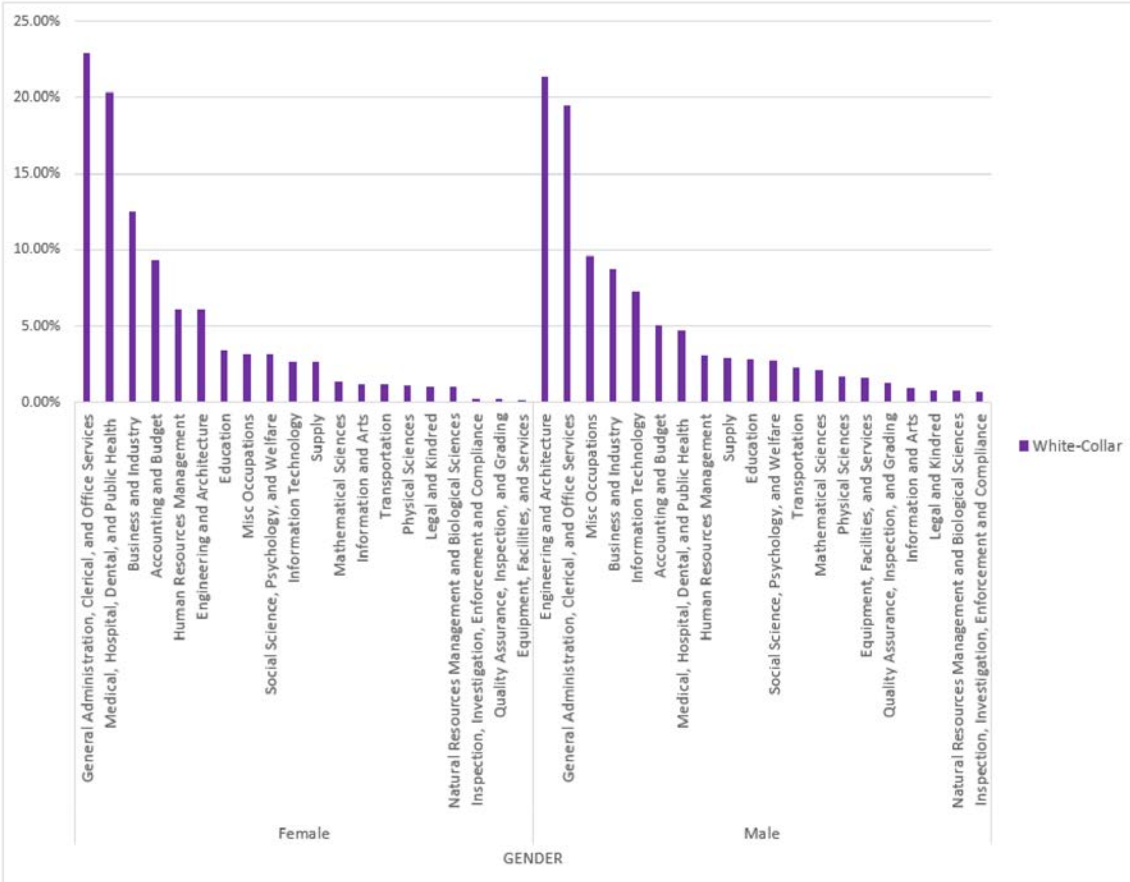



Figure 16. Distribution of Federal Occupational Groups by Gender (White-Collar Only)

### 9. Occupational Families

Occupational families instead of groups categorize blue-collar employees. Thirty-six blue-collar occupational families are "used in classifying trade, craft or labor jobs in the Federal Government" (OPM 2018, p. 3). Figure 17 shows the distribution of occupational families by gender. We find that the majority of both genders belong to the "Warehousing and Stock Handling" occupational family.

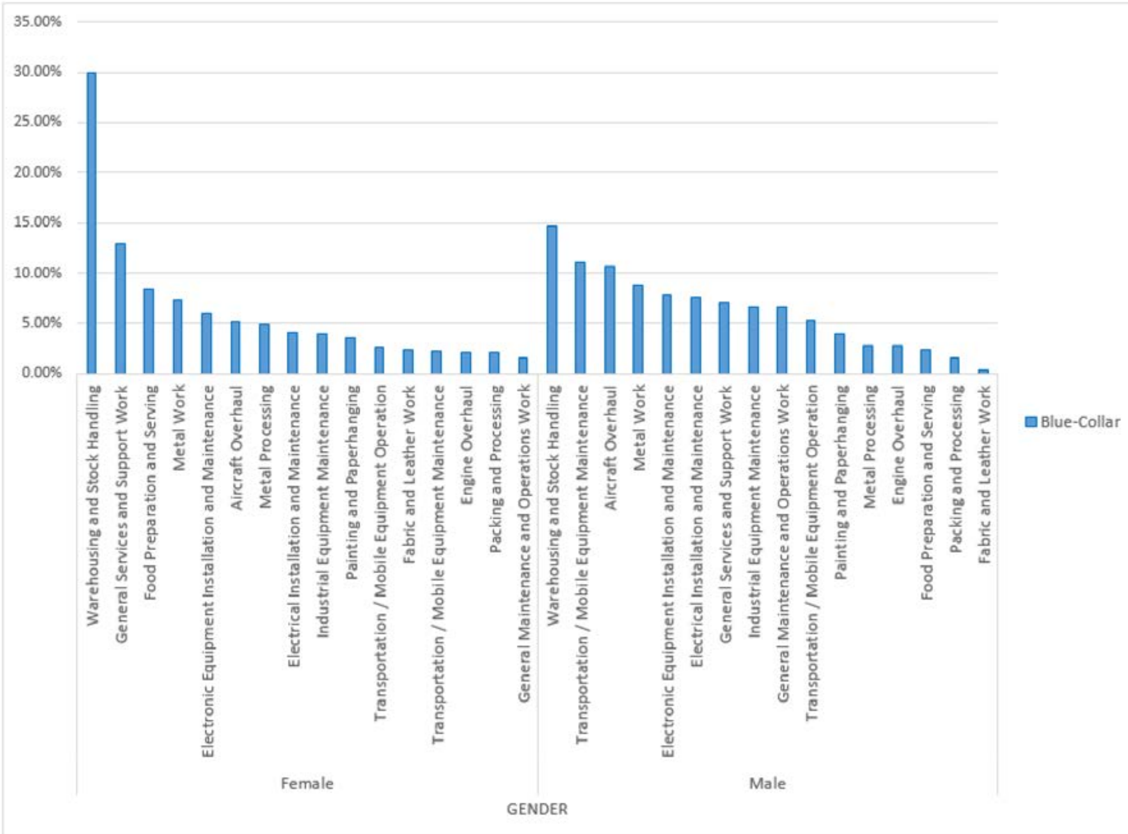



Figure 17. Distribution of Federal Occupational Families by Gender (Blue-Collar Only)

### 10. Life Insurance Coverage

Federal life insurance policy options include “self-only” or “self-plus-family” policies. Another option is to waive the life insurance policy altogether. The civilian personnel data we used in our study does not contain data of marital status or the number of children for civilian employees, so we examine life insurance coverage to gain insights about employees with family members. We also realize that some employees who elect “self-only” or waive life insurance coverage may have spouses or children as well. Figure 18 shows the distribution of life insurance coverage, and we find the majority have “self-only” coverage.

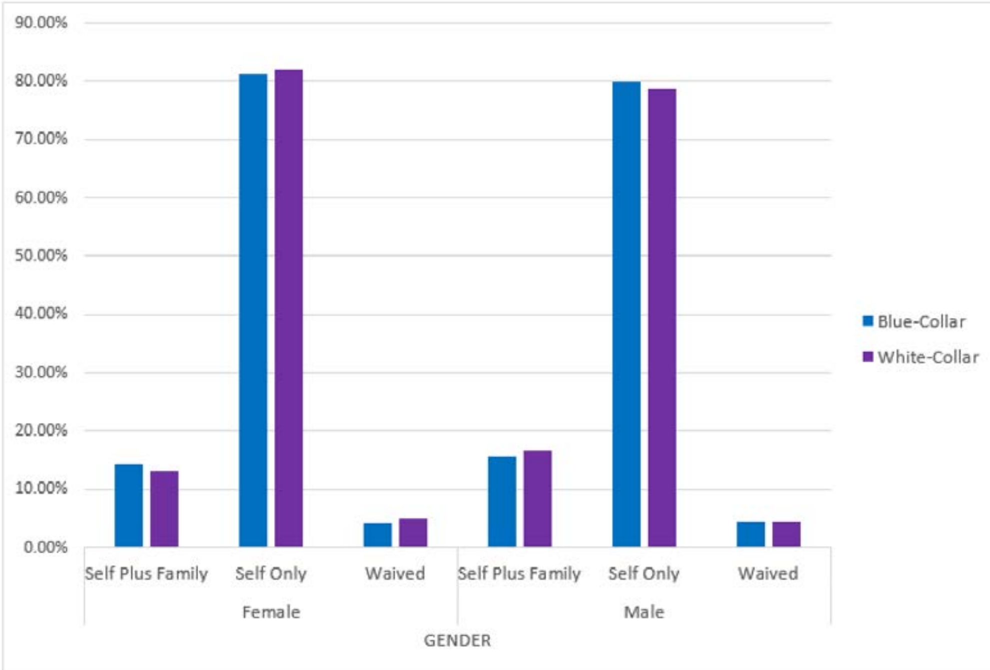



Figure 18. Distribution of Life Insurance Coverage Types by Gender and Collar Type

## B. TIME-VARYING COVARIATES

We now discuss the time-varying covariates of education level, salary, work level, and health insurance. We compare the first snapshot date (date employee begins work) and the last snapshot date (date employee separates or the study period ends) of these covariates to see how much they change throughout the eight-year study period.

### 1. Education Level

The employees have a wide array of education levels. We find that education levels change for more than 25% of the cohort over the course of the study. We see a 10% increase in master's degrees for the white-collar group and also an increase in bachelor's degrees across the entire population, as shown in Figure 19.

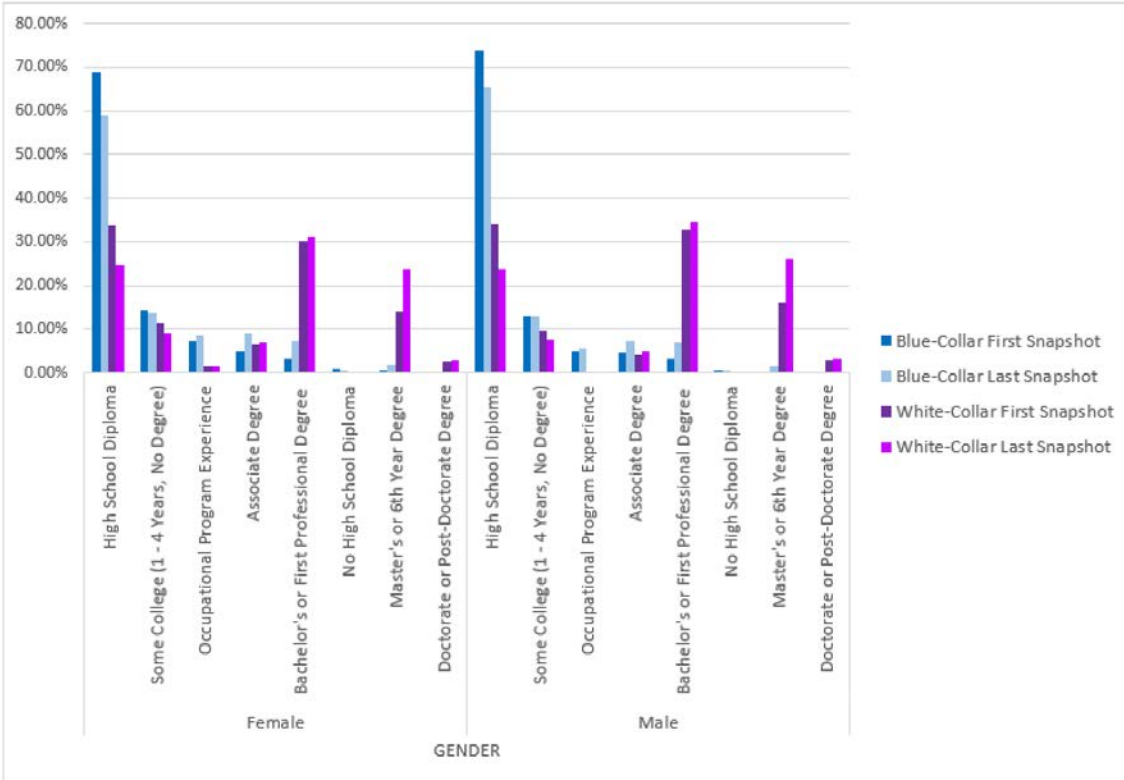



Figure 19. Distribution of Education Levels by First and Last Snapshot, Gender, and Collar Type

### 2. Annual Salary Level

Employee salary levels vary significantly throughout the cohort. We see a dramatic increase in the number of blue-collar employees who earn $50,000 to $80,000 per year, shown in Figure 20. We also see a dramatic rise in the white-collar group of employees who earn more than $80,000 per year.

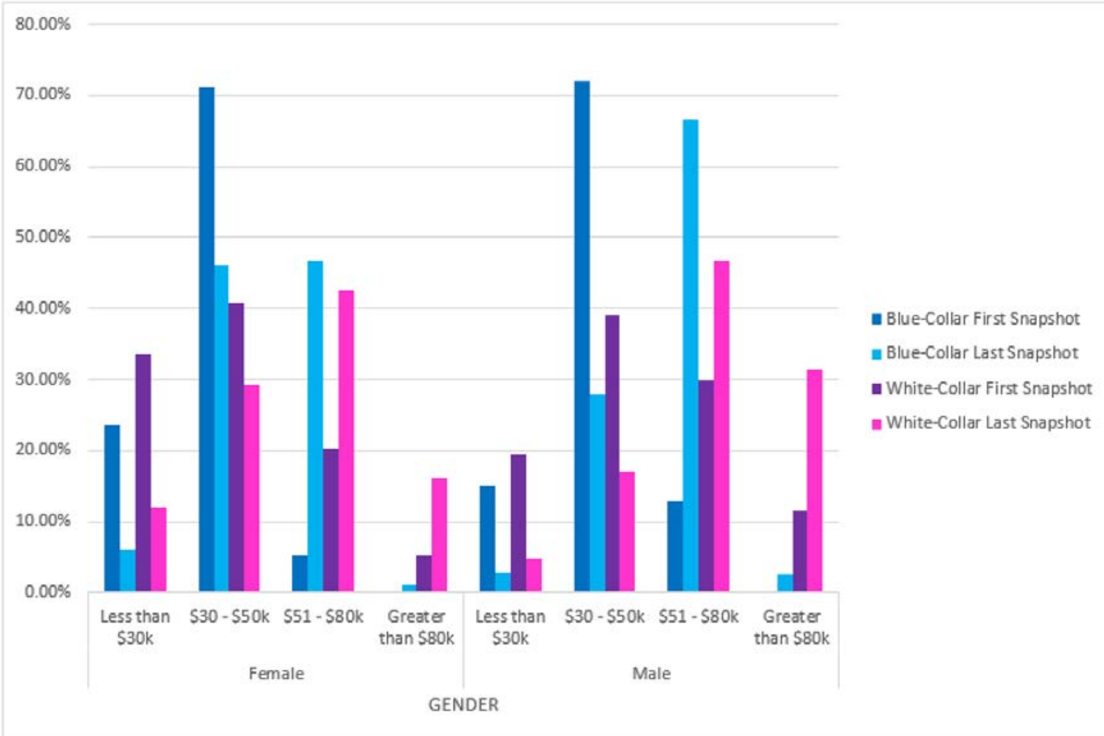



Figure 20. Distribution of Salary Levels by First and Last Snapshot, Gender, and Collar Type

### 3. Work Levels

The employees fall into three work-level categories based on their pay rates. Entry-level, mid-level, and senior-level employees have GS or FWS pay rates of 1–9, 10–12, and 13–15, respectively. We find a significant increase in mid-level and senior-level categories within the female employees as shown in Figure 21. We also see a substantial rise in senior-level blue-collar males.

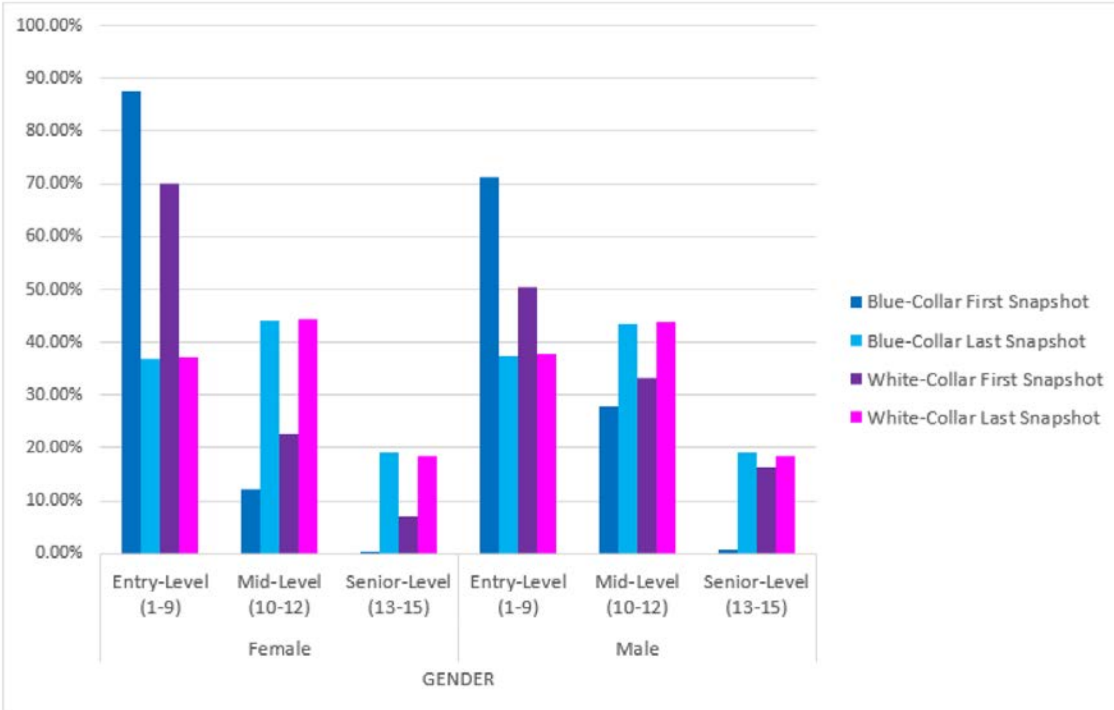



Figure 21. Distribution of Work Levels by First and Last Snapshot, Gender, and Collar Type

### 4. Health Insurance Coverage

We now focus on health insurance coverage to get a sense of which employees have families, based on their coverage. Our understanding is that employees with families may have a lower attrition risk than those without families. We group health insurance coverage into four categories: self-only, self and family, TRICARE For Life (TFL), and declined. "Self-only" coverage implies that an employee has health insurance for themselves alone. We recognize there likely are employees who select "self-only," or who decline health insurance entirely, yet still have families. "Self and family" coverage implies that an employee has at least one dependent family member. "TFL" coverage applies to retired AC servicemembers who receive free healthcare and decline the federal healthcare plan. Employees who decline the healthcare coverage classify as "declined." Figure 22 shows the distribution of healthcare coverage, where we see a noticeable decrease in those employees who decline health insurance and those who receive "self and family" insurance.

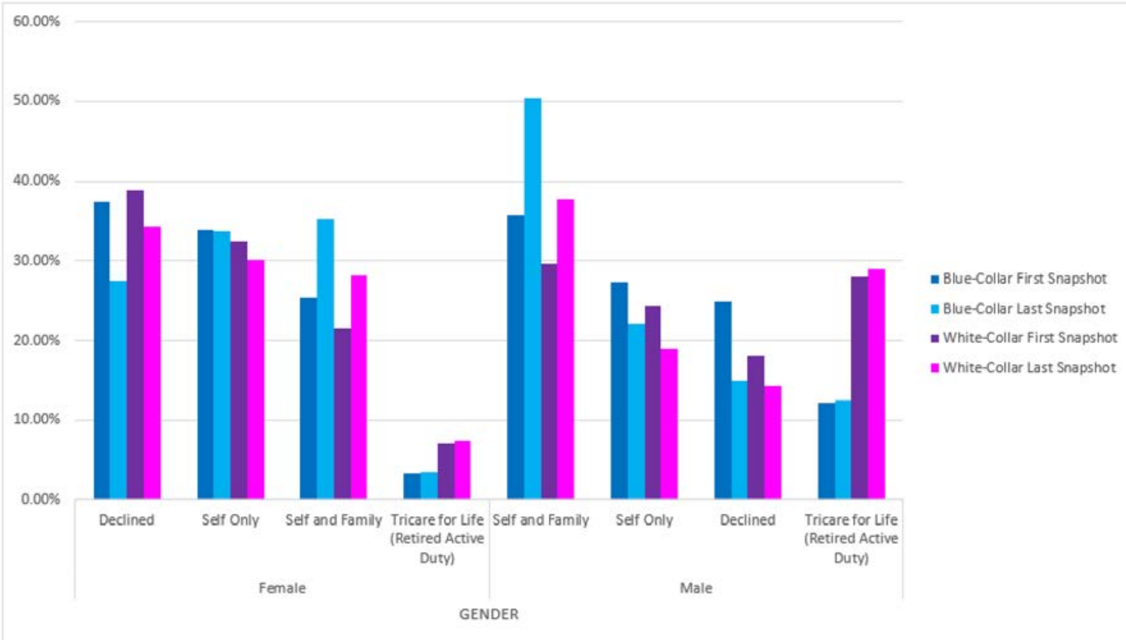



Figure 22. Distribution of Healthcare Insurance Coverage by First and Last Snapshot, Gender, and Collar Type

## C. MILITARY COVARIATES

The covariates in this section pertain to prior AC and prior-or-current RC military servicemembers. We examine these covariates because 44% of the cohort consists of prior AC and prior-or-current RC veterans and we have access to their AC and RC personnel records. This percentage of veterans is higher than the overall 2016 federal government, which consists of approximately one-third veterans (OPM 2017).

### 1. Military Experience

We group employees with or without military experience into three categories: prior AC, prior-or-current RC, and no military experience. The proportions of military experience are similar among the female gender regardless of their collar type. We also find that there are more former AC veterans than RC veterans among collar type and gender. Figure 23 shows the distribution of military experience by gender and collar type of the cohort.

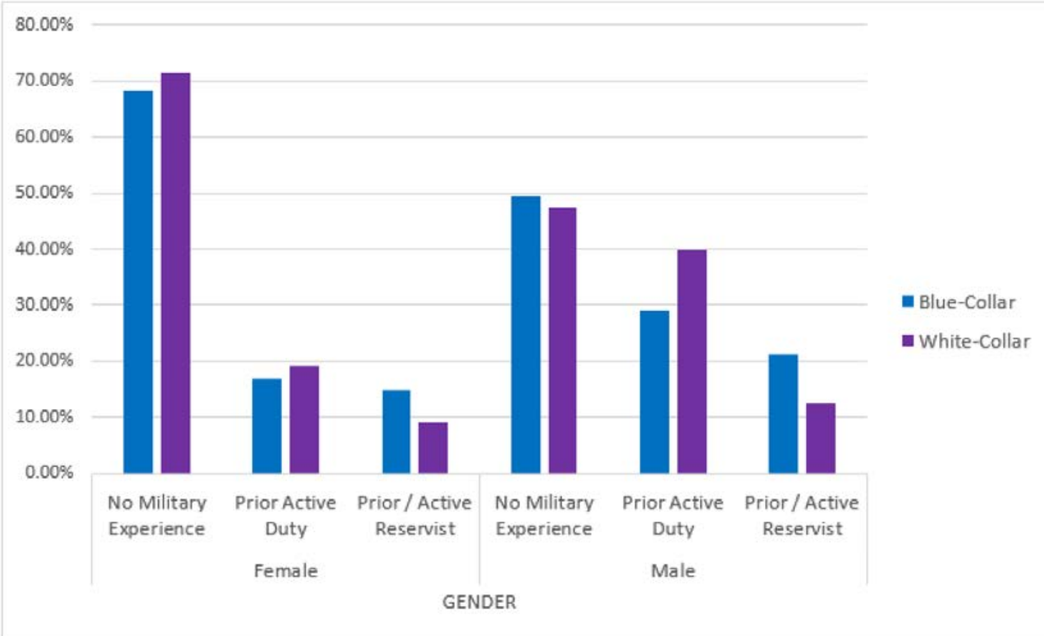



Figure 23. Distribution of Prior Military Service by Gender and Collar Type

### 2. Annuitant Status

An essential characteristic of a prior AC military member is whether or not a member serves 20 or more years to earn a military pension, which may range from $15,000 to $75,000 per year, depending on how long a member serves and their paygrade. Employees with a pension are listed as "annuitants" in their personnel files. We group the annuitants into three categories: retired enlisted servicemember, retired military officer, and no annuitant status. We believe "annuitants" have a much lower attrition rate due to this additional income, which may help to offset low federal salaries. Figure 24 shows the distribution of annuitant categories by gender and collar type. We see that 30% and 13% of white-collar and blue-collar males receive a pension compared to 7% and 3% of white-collar and blue-collar females, respectively.

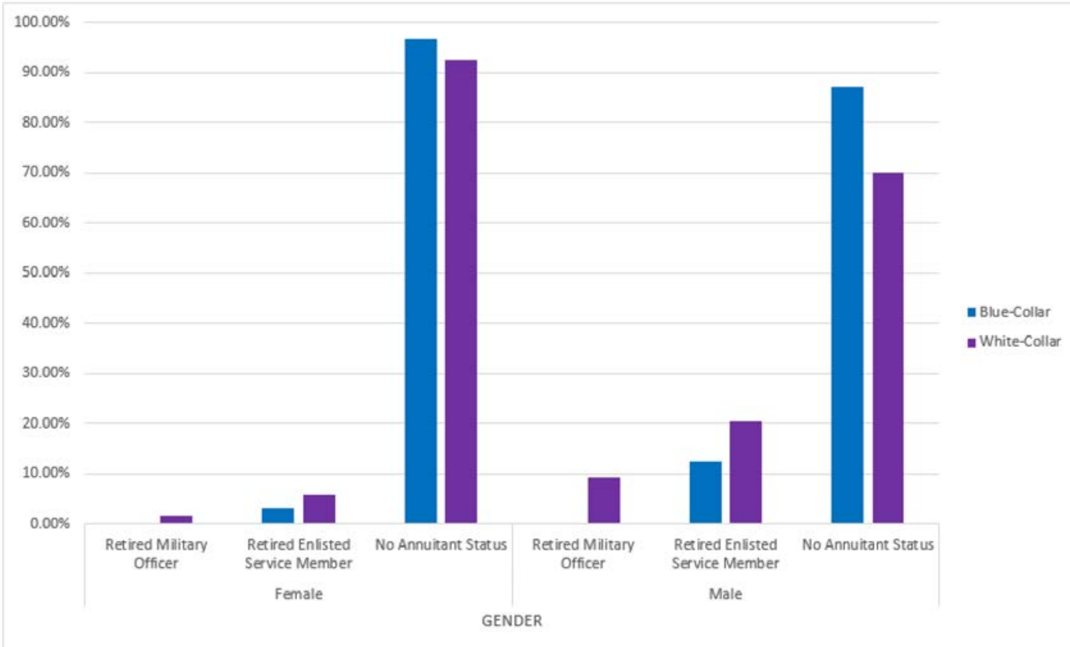



Figure 24. Distribution of Annuitant Status by Gender and Collar Type

### 3. Armed Forces Qualification Test Score Percentiles

We also present the AFQT percentile scores of those with prior military experience to see if there is any correlation between AFQT scores and attrition. We group AFQT scores into five categories: 93–99, 65–92, 50–64, 31–49, and 0–30. These categories reflect the percentile score of an employee on the AFQT. Figure 25 shows the distribution of AFQT scores by gender and collar type. We find that AFQT score proportions are similar between blue-collar and white-collar females. We do, however, see that white-collar males score better overall on the test than blue-collar males.

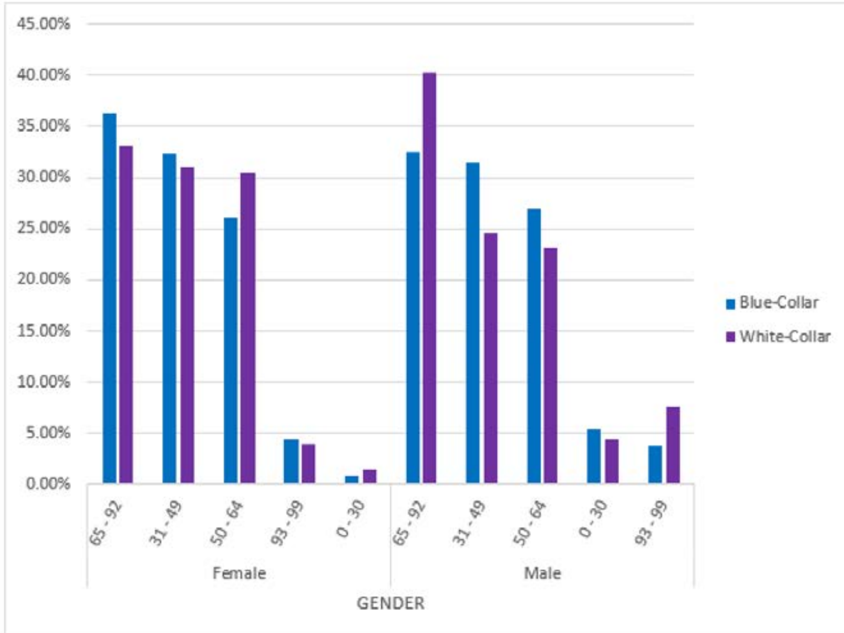



Figure 25. Distribution of AFQT Percentile Scores by Gender and Collar Type

### 4. Military Paygrade Groups

We present the distribution of paygrades of prior AC and prior-or-current RC employees. We group the paygrades into six categories, including E-4 and below, E-5 and above, CWO3 and below, CWO4 and above, O-3 and below, and O-4 and above. The first two categories apply to enlisted servicemembers. The "CWO" categories apply to chief warrant officers and the last two categories apply to military officers. We notice the majority of employees with military experience fit the category of "E-5 and above." Figure 26 shows the distribution of prior AC and prior-or-current RC employee paygrade groups by gender and collar type.

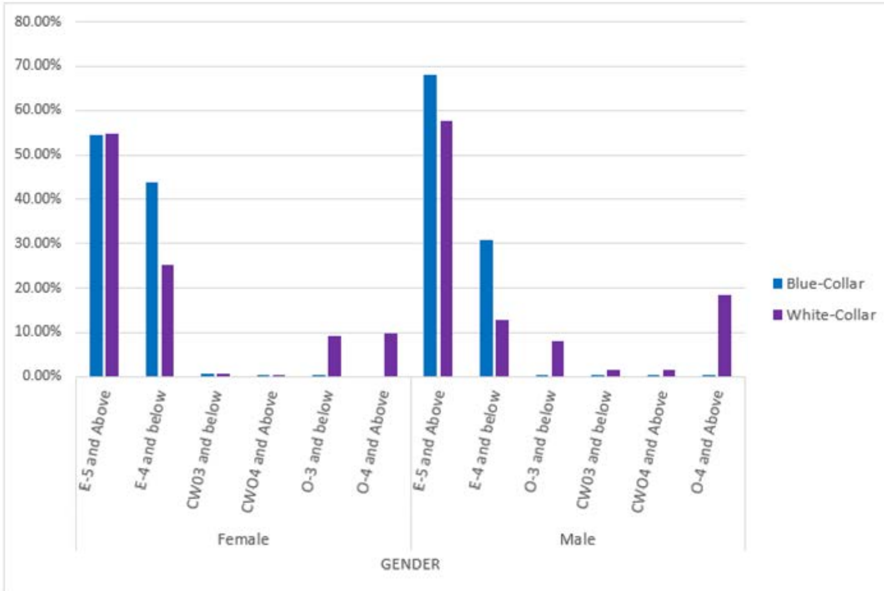



Figure 26. Distribution of Prior Active Duty and Prior-or-Current Reserve Employee Paygrade Groups by Gender and Collar Type

### 5. Active Duty Service Years

Employees with prior AC experience usually have served on active duty from 4 to 20 years or more. We group the years of AC service into five categories: less than 3 years, 3–5 (inclusive) years, 6–10 (inclusive) years, 11–19 (inclusive) years, and 20 years or higher. We are curious if those employees with less than three years of service are the highest attrition risk, due to the employees not completing the standard minimum requirement of four years of service. Figure 27 shows the distribution of years of service by gender and collar type. We see the majority of white-collar males complete 20 years of service. We also find that the majority of blue-collar women complete 3–5 years of service, compared to the white-collar women who achieve 20 years or higher.

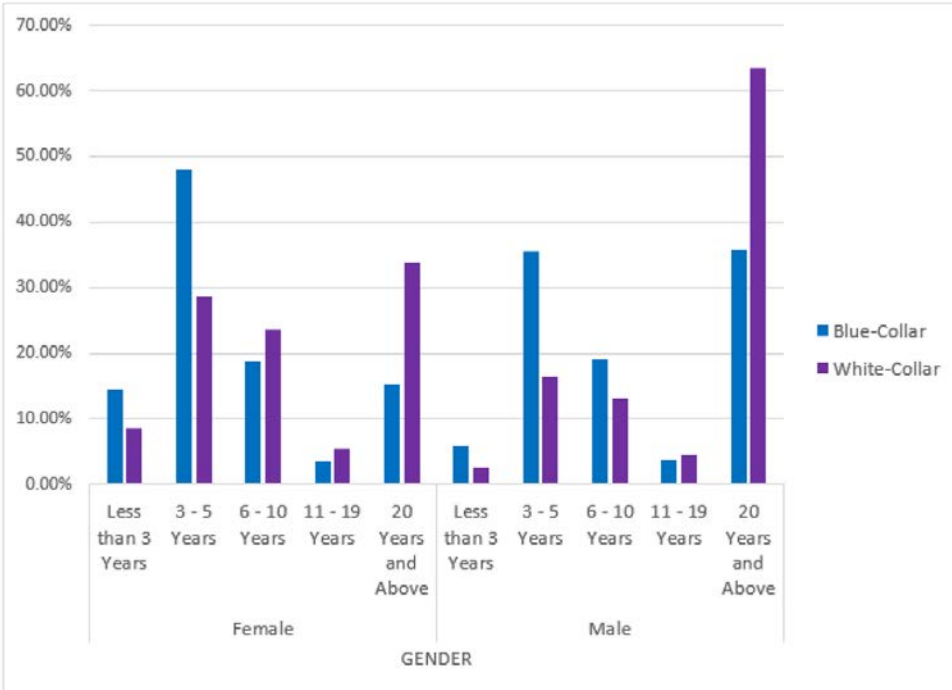



Figure 27. Distribution of Active Duty Service Years by Gender and Collar Type

### 6. Inter-Service Separation Codes

Inter-service separation codes give a detailed description of why an individual is separating from military service (AC or RC). We group these codes into three categories: retirement, expiration of service or other, and red flag behavior. Those employees with separation codes indicating they separate due to retirement are in the "retirement category." Employees who separate from service due to fulfilling their enlistment obligation contract or separate due to any other reason which does not relate to behavior issues or poor performance is in the "expiration of service or other" category. Employees who separate from service due to character or behavior disorder, alcoholism, discreditable incidents, drugs, civil court conviction, fraudulent entry, absence without leave or desertion, discharge in lieu of court-martial, misconduct, pattern of minor disciplinary actions, commission of a serious offense, failure to meet minimum qualifications for retention, and unsatisfactory performance fall under the "red flag behavior" category. Figure 28 presents the distribution of inter-service separation codes by gender and collar type.

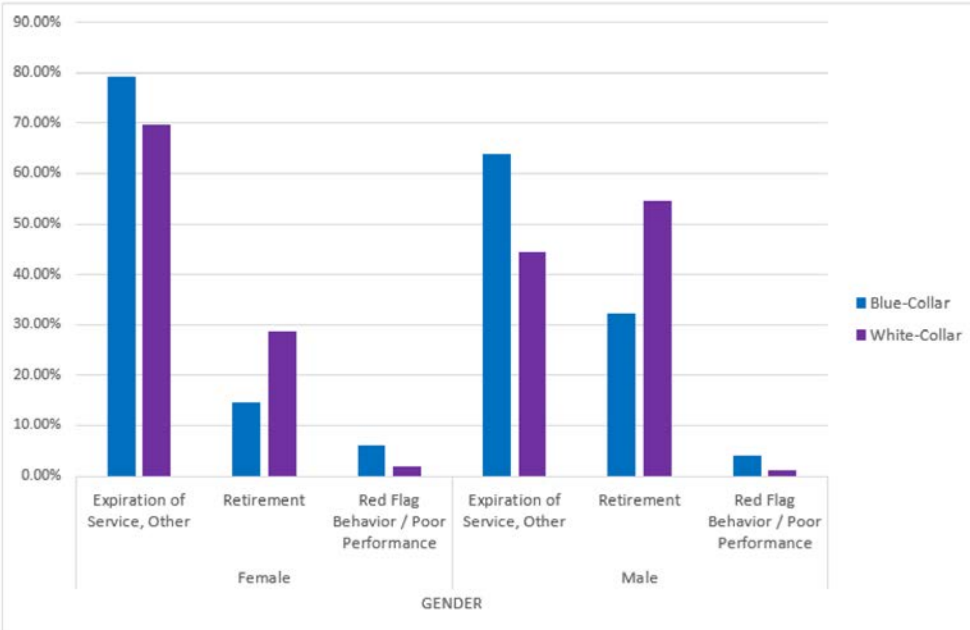



Figure 28. Distribution of Inter-Service Separation Codes by Gender and Collar Type

# IV. SURVIVAL ANALYSIS MODELING AND EVALUATION

In this section, we present an introduction to survival analysis modeling and an evaluation of the survival analysis findings. Many of the plots described in this section can be found in Appendix B.

## A. SURVIVAL ANALYSIS MODELING

The fourth phase in the CRISP-DM process is "modeling" in which, "various modeling techniques are selected and applied" (Wirth and Hipp 2000, p. 6). The approach selected for our study is survival analysis. "Survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs" (Kleinbaum and Klein 2005, p. 5). "Time" in our case refers to the number of years from the date which an employee begins DoD employment until the "event" of attrition from DoD employment is reached. Active component (AC) servicemember attrition studies that focus on a short time period of behavior have been widely conducted in the past. Buttrey et al. (2018) note that for these studies, where the response variable is attrite or not attrite, for a short time period and where the entire cohort starts at the same time, classification methods such as logistic regression work well. Due to the diversity of the DoD workforce and the fact that the employees do not join for a specified period of time, survival analysis is a good approach to forecast attrition behavior.

### 1. Survival Function

Survival analyses are used to estimate a survival function, $S(t)$, which represents the probability that a DoD employee survives longer than $t$ units of time. For our study, $t$ is the number of years of employment beginning in 2009. The survival function is the foundation of survival analysis because the survival probabilities provide a wealth of information for periods of time during a study (Kleinbaum and Klein 2018). The survival function is non-decreasing with $S(0)$ defined to be one, indicating that all employees are employed on the date they are hired ($t = 0$). For this study, the survival function is estimated based on data from the eight-year study period, thus $S(t)$ can only be estimated for $t$ less than eight years. We plot all our survival functions using the R environment (R Core Team 2013) and the R package "survminer" (Kassambara and Kosinski 2017).

### 2. Kaplan-Meier Estimate

We use the KM estimator of Kaplan and Meier (1958), a nonparametric survival function estimator. Survival analysis based on the KM estimator is commonly used to study survival times or times until major events of people and other living organisms (e.g., see Kleinbaum and Klein 2018 for further discussion).

The KM estimator accommodates both left-truncated and right-censored data. Since all employees in the dataset are new appointments, we observe their career paths starting at time $t = 0$. Thus, none of the records in the dataset are left-truncated. However, all employees who are still employed by DoD at the end of the eight-year study are considered right-censored. For these employees, we do not observe their attrition date, but we do know their censoring time, i.e., the time (since appointment) that they are lost to the study. Even though the end of the study period (2017-03-31) is the same for all employees, censoring times vary between seven and eight years depending on the employees' 2009 appointment dates.

The KM estimator is constructed by expressing a survival probability as the product of conditional probabilities, each of which is easily estimated. Table 3 illustrates the computations involved. The KM estimator is a non-decreasing step function with steps only at observed attrition time $t_0 = 0 < t_1 < t_2 < \ldots$ The number of employees "at risk" at the start of an interval is the total number in the study minus the number who have attrited or who have been lost due to censoring in earlier intervals. Thus, the estimated probability of surviving an interval among those at risk is the ratio of the number surviving to the number at risk, and the estimate of the unconditional probability is the product of the conditional ones. Two features of the KM estimator are that it has an undefined right tail when observations are right-censored at the end of the study and that with large numbers of observations; the KM estimator appears fairly smooth as in Figure 1. With smaller numbers of observations, the steps in the KM estimator become more apparent.

Table 3. Example of KM Estimator Method

| Time Period $[t_{i-1}, t_i)$ | Number of Employees at risk at $t_{i-1}$ | Number of Employees who attrite at $t_{i-1}$ | Number of Employees censored at $t_i$ | KM Estimate: Estimate of $S(t)$ for $t_{i-1} < t < t_i$ |
|---|---|---|---|---|
| $[0, t_1)$ | 100 | 0 | 0 | $\frac{100}{100} = 1.0$ |
| $[t_1, t_2)$ | 100 | 5 | 0 | $\frac{95}{100} * \frac{100}{100} = 0.95$ |
| $[t_2, t_3)$ | 95 | 10 | 0 | $\frac{85}{95} * \frac{95}{100} * \frac{100}{100} = 0.85$ |
| $[t_3, t_4)$ | 85 | 10 | 5 | $\frac{75}{85} * \frac{85}{95} * \frac{95}{100} * \frac{100}{100} = 0.75$ |
| $[t_4, t_5)$ | 70 | 20 | 50 | $\frac{50}{70} * \frac{75}{85} * \frac{85}{95} * \frac{95}{100} * \frac{100}{100} = 0.54$ |
| $[t_5, \infty)$ | 0 | 0 | 0 | Undefined |

### 3. Survival Trees

We also fit survival trees where the KM estimator is used to estimate the survival functions for subsets of the data to gain further insights. Our survival trees use the algorithms of Hothorn and Zeileis (2015), contained in the R package “partykit.” We also use the R package of “LTRCtrees” constructed by Fu and Simonoff (2017) to account for right-censored data. Survival trees are very similar to regression and classification trees (Breiman et al. 1984). At each step, the survival tree algorithm splits the data into two subsets or nodes using one of the covariates. The covariate and splitting criteria are chosen (among all covariates and splitting criteria) to be the ones with the smallest p-value for the null hypothesis that the two subsets have the same survival function. We only fit the time-constant covariates to our survival trees due in part to PDE computational limitations involved in trying to fit multiple time-varying covariates to a survival tree model.

## B. EVALUATION OF FINDINGS

The next phase of the CRISP-DM process is "evaluation" in which, we evaluate the results of our survival analysis model. This section covers the findings from the exploratory analysis of all the covariates, and several of the covariate survival analysis plots referenced in this section are found in Appendix B.

The analysis of the covariates is exploratory, so the 95% confidence intervals displayed are merely an indication of the relative size of the particular subset. We also do not test hypotheses during the exploratory analysis. Additionally, we acknowledge that there are other confounding variables in the analysis which are not accounted for.

### 1. Gender

The PPS (2014) find that women in the federal government have a higher attrition rate than their male counterparts. An evaluation of the survival function for DoD blue-collar and white-collar males and females shown in Figure 29 shows a similar trend. We find that the blue-collar and white-collar females both have lower survival probabilities than their male counterparts. We also see that blue-collar males and females have higher survival probabilities than white-collar men and women, respectively.

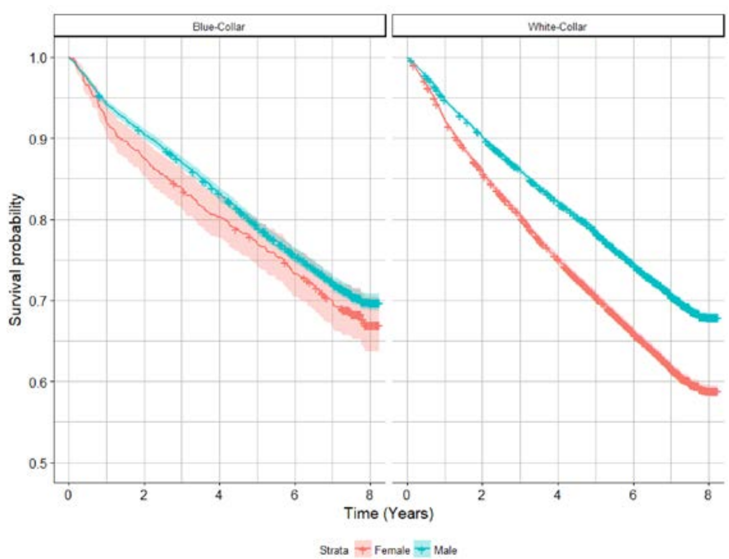



Figure 29. KM Estimated Survival Functions by Gender and Collar Type

We also consider the amount of time it takes for an employee to leave their job, by examining the distribution of time an employee works prior to separating. According to Light's report (2011), "while overall attrition remained low, nearly 25% of new federal employees leave within two years, with the rate as high as 30% in certain departments" (p. 38). Figure 30 shows that this is also true for DoD employees. The attrition rate is greatest in the first two years among both genders and collar types. We see attrition rates between years two through seven are between 10% and 15% and attrition is the lowest in the last year of the study.

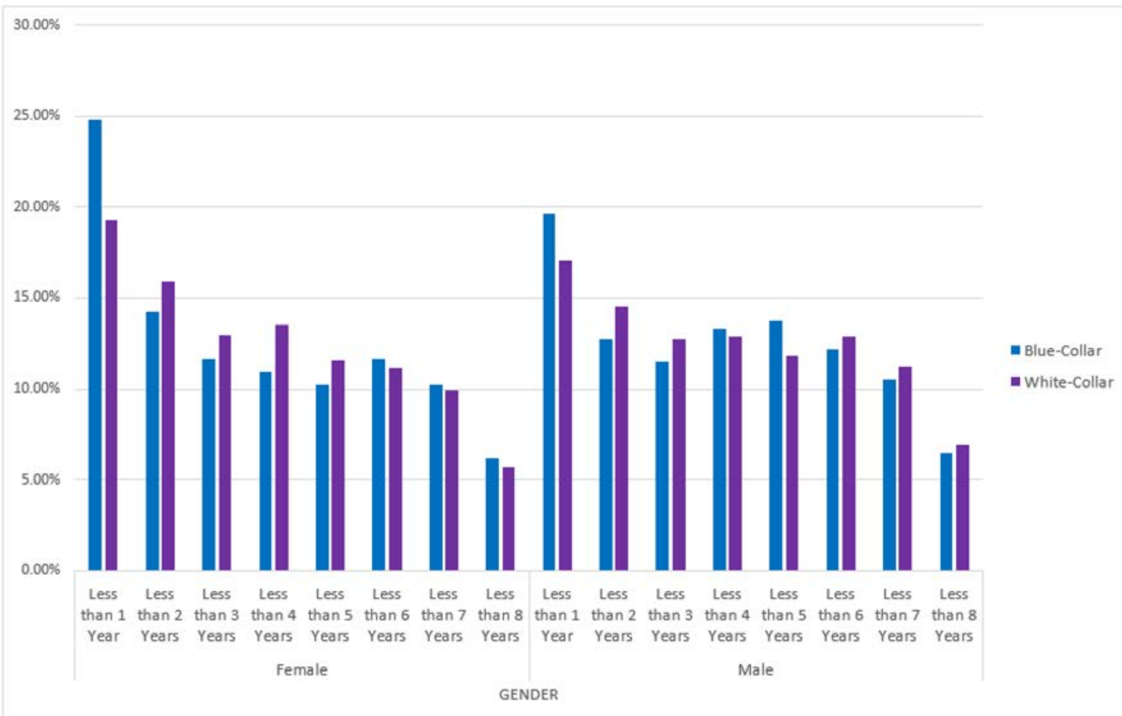



Figure 30. KM Estimated Probability of Attrition by Gender and Collar Type

### 2. Age

The 2009 cohort is separated into 11 age groups to compare the distribution of survival probabilities. Employees less than 20 years of age and those of age 55 or greater have the highest attrition probabilities among the white-collar cohort, as shown in Figure 31. We also find that white-collar females have a higher attrition probability than white-collar males across all age groups.

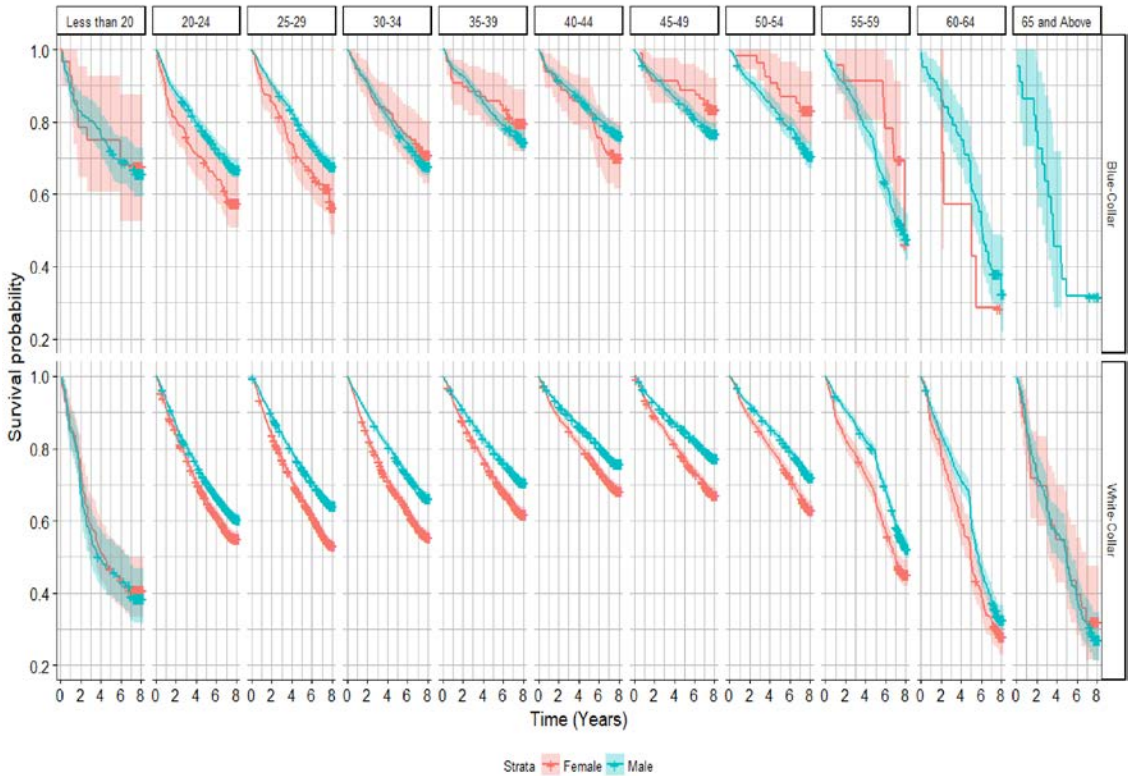



Figure 31. KM Estimated Survival Functions by Gender, Age Group, and Collar Type

### 3. Occupational Groups, Categories, and Families

PPS (2014) find that the top two highest attrition occupational groups were the administration, operations, and general management group (code 0300), and the medical, dental, and public health group (code 0600). Similar to the PPS (2014) study, Figure 32 shows that the medical, dental, and public health group has the lowest survival probability among the white-collar occupational groups at approximately 50% for males and 43% for women. We find that the administration, operations, and general management group have nearly 72% and 63% eight-year survival probabilities for males and females, respectively. As can be seen in Figure 33, we find that the white-collar occupational category with the highest attrition rate is "clerical" among both genders and collar types. We find that "administrative" has the lowest rate of attrition.

In Figure B1, we find that the food preparation and services (code 7400) and the ammunition, explosives, and toxic materials (code 6500) blue-collar occupational families have the lowest survival probabilities among both genders. The tables of occupational groups and family codes can be found in Appendix C.

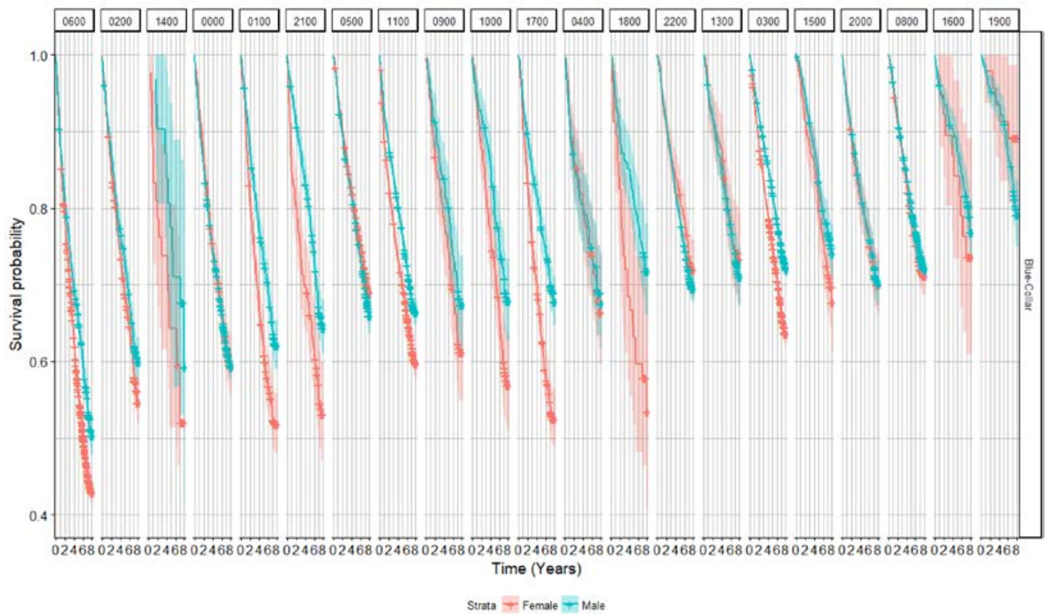



Figure 32. KM Estimated Survival Functions by Gender, Occupational Group, and Collar Type

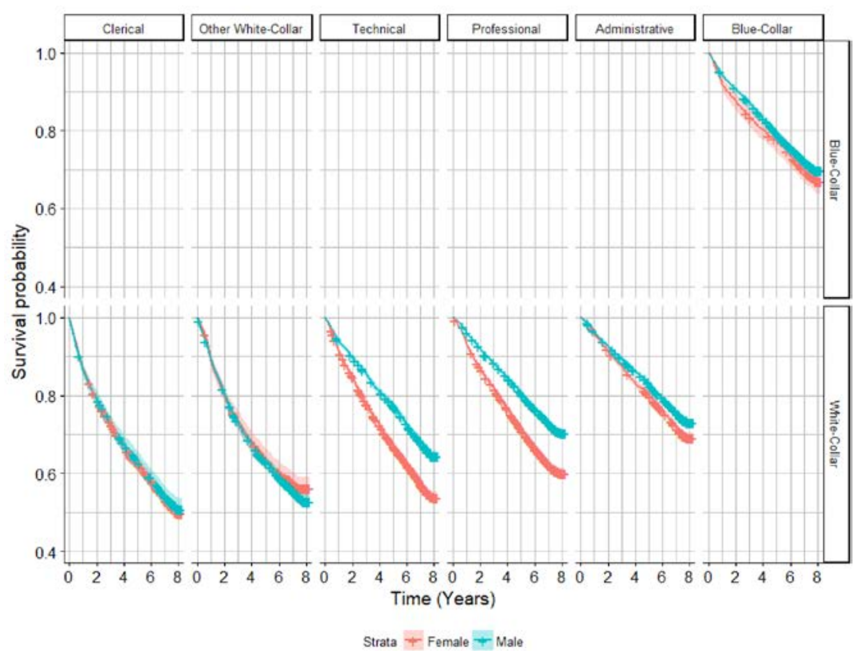



Figure 33. KM Estimated Survival Functions by Occupational Category, Gender, and Collar Type

### 4. Health and Life Insurance Coverage

Health insurance coverage is examined to analyze a category of employees who have a family. We also categorize employees who are retired AC military servicemembers and have TFL health insurance coverage, because the data shows these members "declined" the federal health insurance coverage options. We find in Figure 34 that employees with family members have a higher survival probability than those who decline the federal health insurance or have "self-only" coverage. This finding stresses the importance of obtaining marital status and the number of family members in the household information for future DoD employee attrition studies. We also see that the retired AC males have the highest survival probability.

We also find similar results with KM estimated survival function of life insurance coverage. Figure B2 shows that employees, regardless of gender or collar type, with family life insurance coverage have a higher survival probability than those who waive the coverage or have "self-only" coverage. We find that individuals who waive life insurance coverage have the lowest survival probability among the three categories.

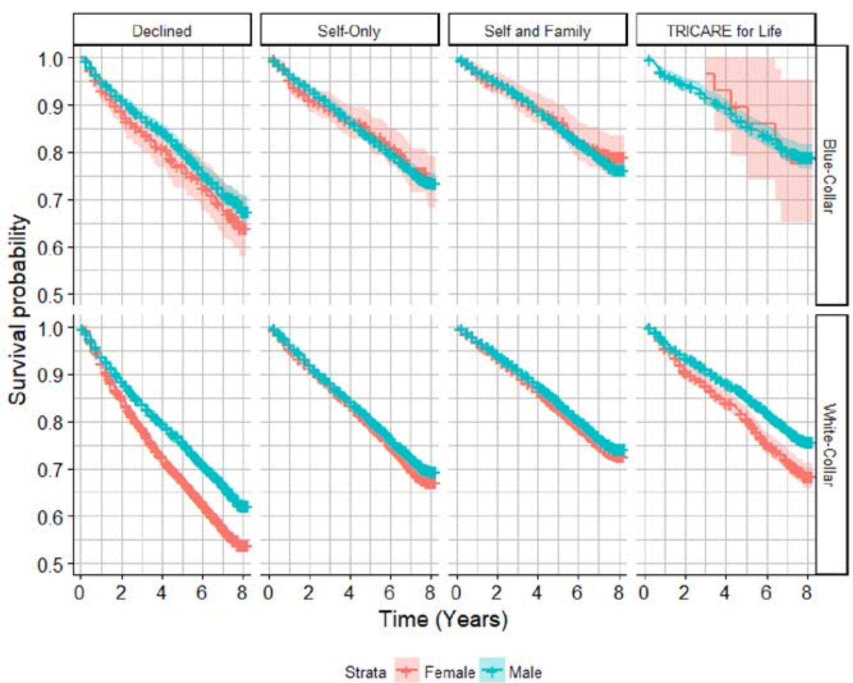



Figure 34. KM Estimated Survival Functions by Gender, Health Insurance Coverage, and Collar Type

### 5. Annual Salary and Work Level

Annual salaries are divided into entry-level (less than $30,000), middle ($30,000-to-$50,000), upper middle ($50,000-$80,000), and high (greater than $80,000) income classifications. Figure 35 shows that annual income is directly related to attrition and the higher the annual salary, the higher the employee survival probability among both genders and collar types.

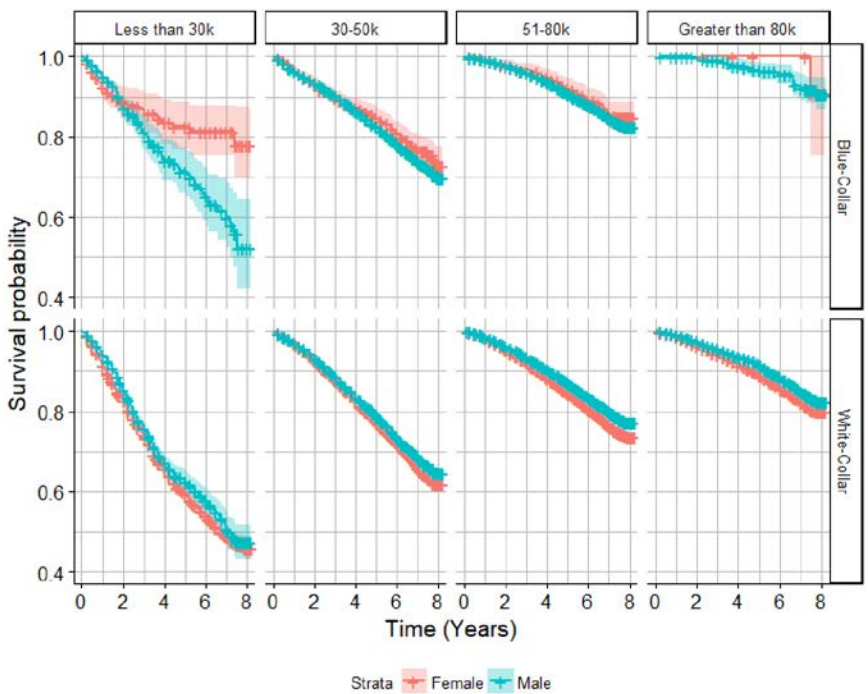



Figure 35. KM Estimated Survival Functions by Gender, Annual Salary, and Collar Type

The PPS (2014) find that federal government entry-level employees had the highest amount of attrition compared to mid-level and senior-level employees. Figure 36 shows that this also holds for DoD employees of both gender and collar types. We see white-collar male and female entry-level employees have only a 60% and 53% survival probability, respectively.

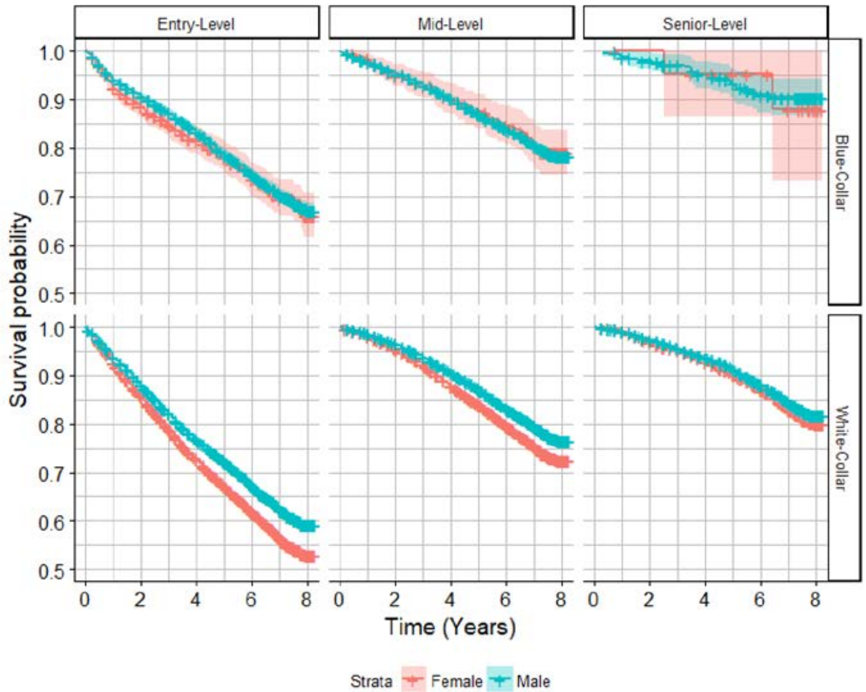



Figure 36. KM Estimated Survival Functions by Gender, Work Level, and Collar Type

### 6. Education

Education is divided into seven categories based on the education level of the employee. In Figure 37, we find that white-collar male employees with a doctorate have the lowest survival probability among white-collar males. We see that white-collar males and females with an associate degree, bachelor's degree, or master's degree have a higher survival probability than those employees without a college degree. It is hard to compare education attributes among the blue-collar cohort, due to the majority of the cohort possessing only a high school diploma.

### 7. Annuitant Status

Annuitant status is divided into five categories. We differentiate between those prior AC and reserve component (RC) employees who either did serve in the military 20 years or longer to gain annuity status or did not (denoted by "none"). Figure 38 shows that retired enlisted servicemembers have the highest survival probability. We also find that prior AC and RC employees who do not have annuitant status have a lower survival probability than employees with no previous military experience.

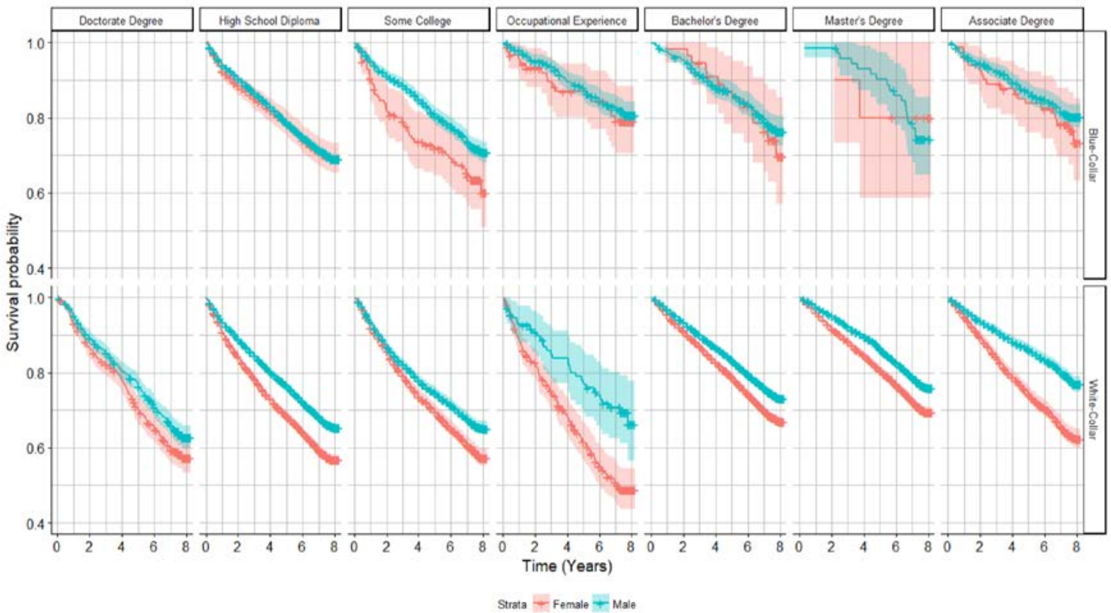



Figure 37. KM Estimated Survival Functions by Gender, Education Level, and Collar Type

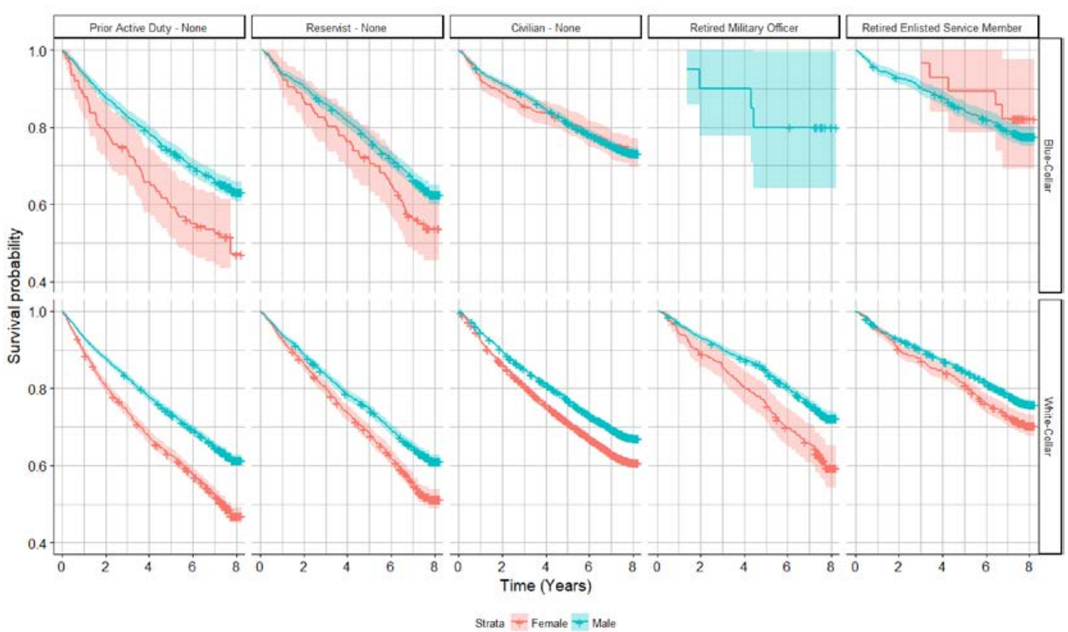



Figure 38. KM Estimated Survival Functions by Gender, Annuitant Status, and Collar Type

### 8. Years of Active Duty Military Service, Military Paygrades, and Prior Federal Creditable Service Years

Years of AC military service are divided into five categories based on the number of years an employee was on active duty. Figure 39 shows that an employee's number of years of AC service is directly related to attrition and the higher the number of years, the higher the employee survival probability among both genders and collar types. We also find similar results when we examine the KM estimated survival functions of the paygrades of prior AC and RC in Figure B3. The survival probability is higher among senior ranking enlisted servicemembers and military officers, who have more years of service than junior enlisted servicemembers and military officers.

Figure B4 shows the KM estimated survival functions for prior FCS years. We find that employees with no previous FCS years have the highest survival probability and those with 20 or more FCS service years have the lowest among both gender and collar types.

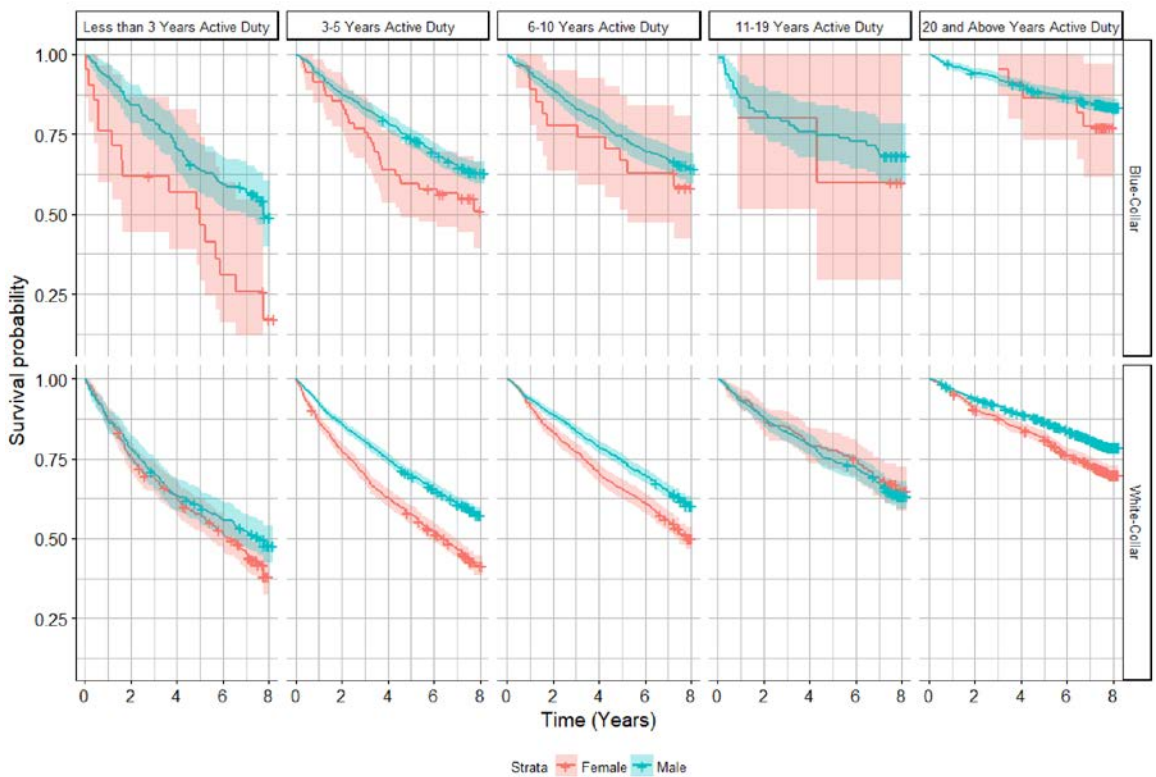



Figure 39. KM Estimated Survival Functions by Gender, Years of Active Duty Military Service

### 9. Race-Ethnicity, Branch of Service, and Bureau of the Census Division

Figure B5 shows the KM estimated survival functions of race-ethnicity by gender and collar type. We find that the Asian and Pacific Islanders have the highest survival probabilities among both gender and collar type. We also see that the American Indian or Native Alaskan have the lowest survival probabilities, but the sample size for this race-ethnicity is relatively small.

The PPS (2014) found that the Army had the highest attrition among all government agencies. We see in Figure B6 that this is also true for the DoD. Employees working for the Army have the lowest survival probabilities, while those working for the Navy have the highest. We also see in Figure B7 that the Bureau of the Census "mountain" region has the lowest survival probability for white-collar males, which is most likely attributed to the large population of Army bases in that region. We also find that "east north" and "east south" regions have the highest survival probability among both genders and collar types.

### 10. Military Experience, AFQT Percentile, and Inter-Service Separation Codes

Approximately 44% of the 2009 cohort has either AC or RC military experience. We find in Figure B8 that RC personnel have the lowest survival probability among both genders and collar types. We also find that prior AC males have the highest survival probability among white-collar employees, but male blue-collar employees with no military experience have a higher survival probability than blue-collar males with prior AC military experience. In Figure B9, we see that AFQT scores do not provide much insight into survival probabilities, but we do see that employees with the lowest AFQT scores have the highest survival probabilities, but the sample size for this group is relatively small. In Figure 39, we find that the employees who were discharged for "red-flag behavior" have the lowest survival probability, but the sample size was only 340 employees, so this finding should seek further analysis with a larger sample size. We also see employees who separate from active duty due to "retirement" have the highest survival probability among both genders and collar types.

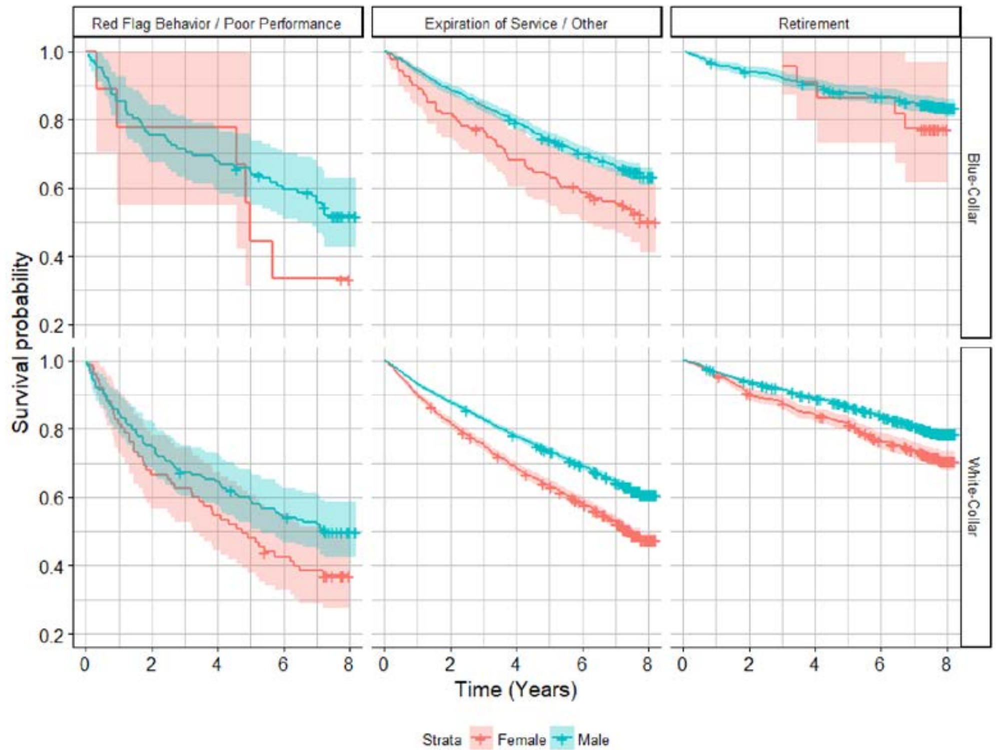



Figure 40. KM Estimated Survival Functions by Gender, Inter-Service Separation Reason, and Collar Type

### 11. Survival Trees

We fit a survival tree with 11 time-constant covariates including gender, age, annuitant status, census bureau division, occupational category, life insurance coverage, branch of service, prior FCS years, military experience, race-ethnicity, and collar type. We prune the tree by setting the p-value of the splits to less than 0.001 to find the most relevant covariates. The most relevant covariates based on the tree splits are occupational category, annuitant status, census bureau division, branch of service, military experience, and gender as shown in Figure 41. The survival tree splits confirm our previous findings based on the differences in survival probabilities found for these covariates. The abbreviation "Inf," found at each terminal node, means that the median time to attrition cannot be estimated from the KM estimates, because the median is greater than eight years, the length of the study period. One terminal node, node 5, has a median time to attrition of seven years. A detailed printout of the pruned and unpruned survival trees is shown in Tables B1 and B2, respectively.

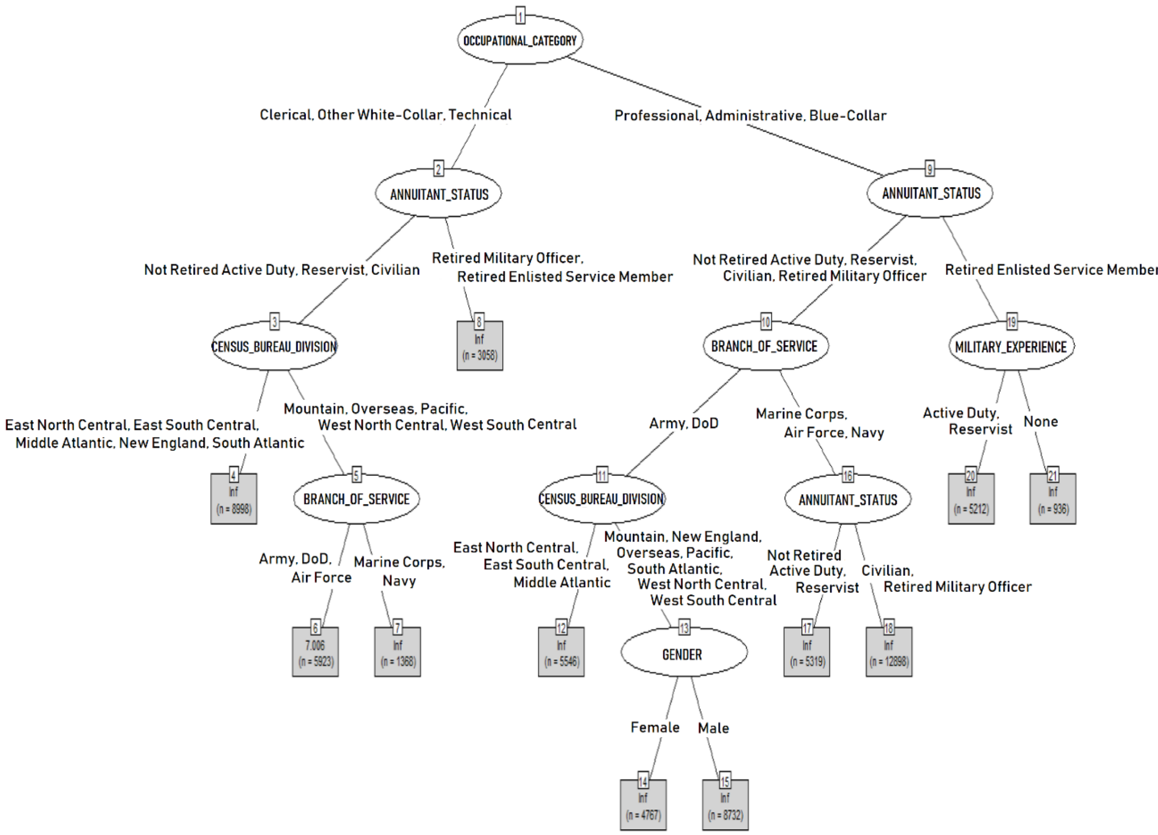



Figure 41. Survival Tree Fits for Gender, Age, Annuitant Status, Census Bureau Division, Occupational Category, Life Insurance, Branch of Service, Prior FCS Years, Military Experience, Race-Ethnicity, and Collar Type

We also plot the survival functions for all 84 terminal nodes of the unpruned tree as shown in Figure 42. This allows us to answer more specific questions about the cohort. For example, the red survival function represents the approximately 60% eight-year survival probability of 44-year-old females with the following characteristics: employed by the DoD, occupational category of “professional,” and no military experience.

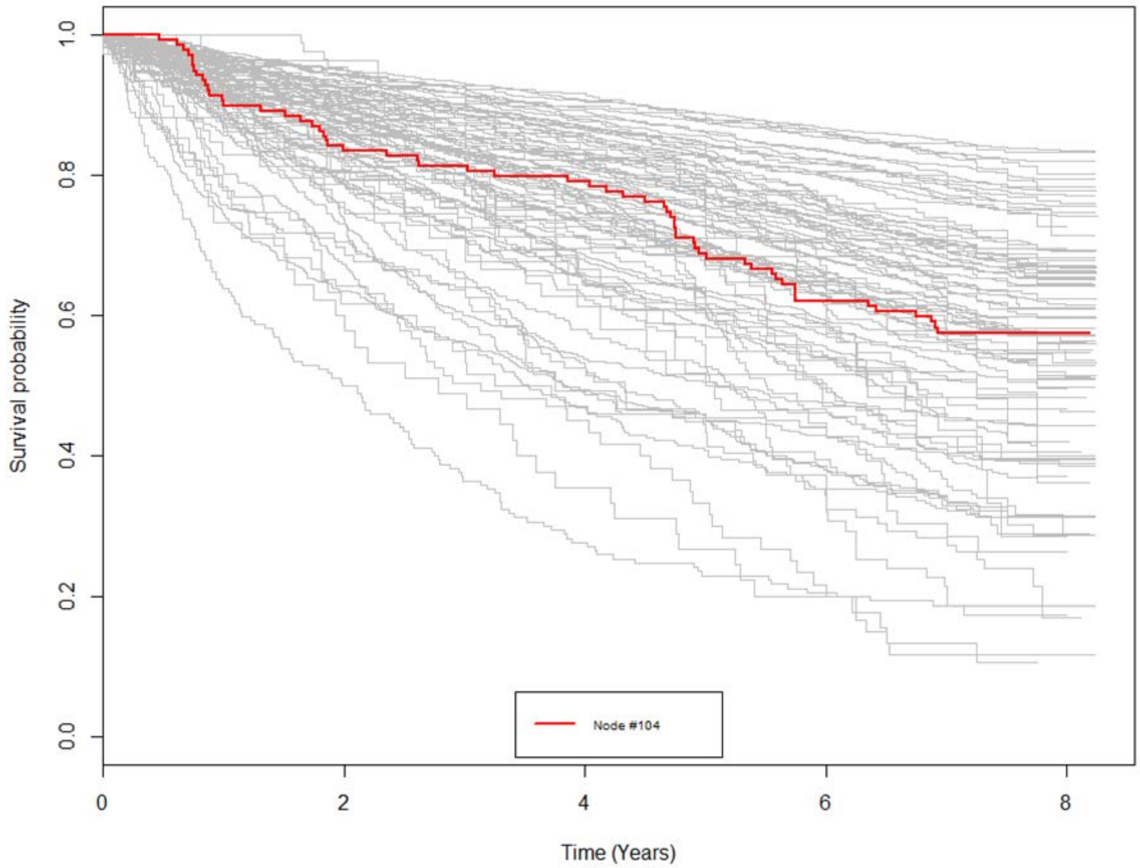



Figure 42. KM Survival Function Estimates for Females, Age 44, "DoD" Branch of Service, "Professional" Occupational Category, and No Military Experience (Node 104)

We also fit an additional survival tree for the employees with prior military experience (prior AC and prior-or-current RC). The tree consists of six time-constant covariates including gender, age, military experience, AFQT test score percentiles, military paygrade groups, and military service separation reasons. We also prune the tree by setting the p-value of the splits to less than 0.001 to find the most relevant covariates. We find that the most relevant covariates are military service separation reasons and gender as shown in Figure 43. The survival tree splits also confirm our previous findings based on the differences in survival probabilities found for these covariates. A detailed printout of the unpruned survival tree is presented in Table B3.

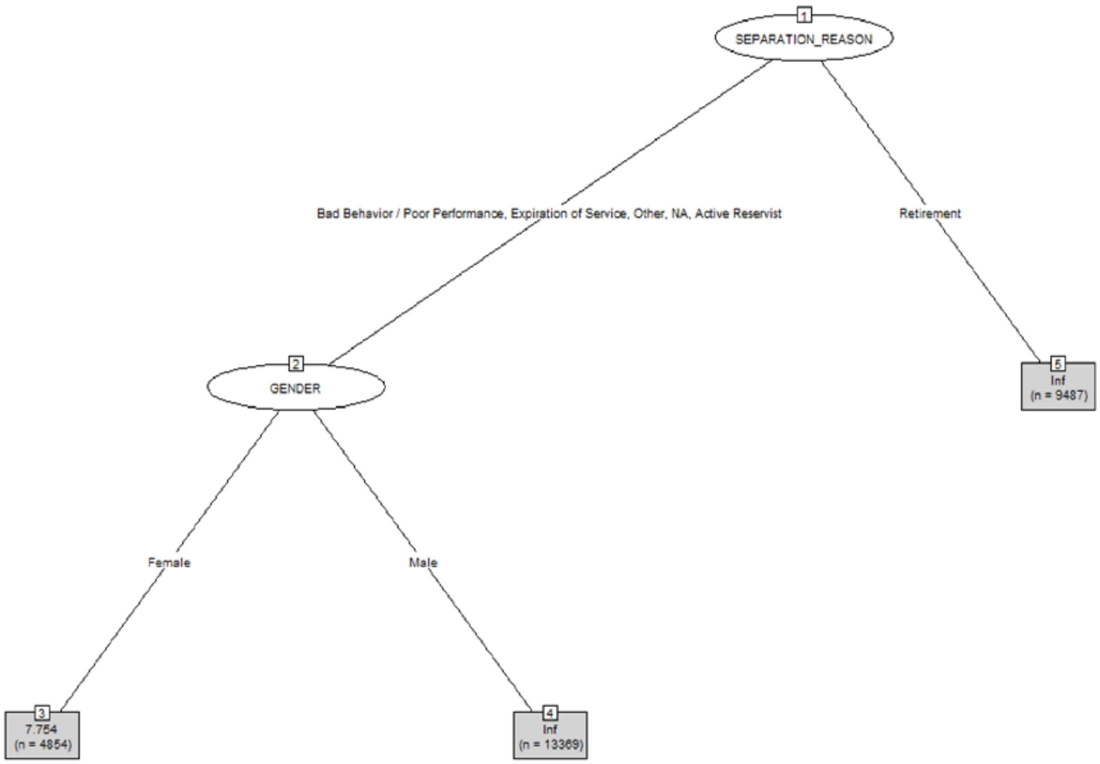



Figure 43. Survival tree fits for Gender, Age, Military Experience, AFQT Percentile Scores, Military Paygrade Groups, and Military Service Separation Reason

We also plot the survival functions for all 23 terminal nodes as shown in Figure 44. As an example, we use these survival functions to compare the survival probabilities between 44-year-old male and female military retirees. We find that male retirees (represented by the blue survival function) have approximately 80% survival probability compared to 70% for female retirees (represented by the red survival function). Terminal Nodes 39 and 42 are the genders of female and male, respectively.

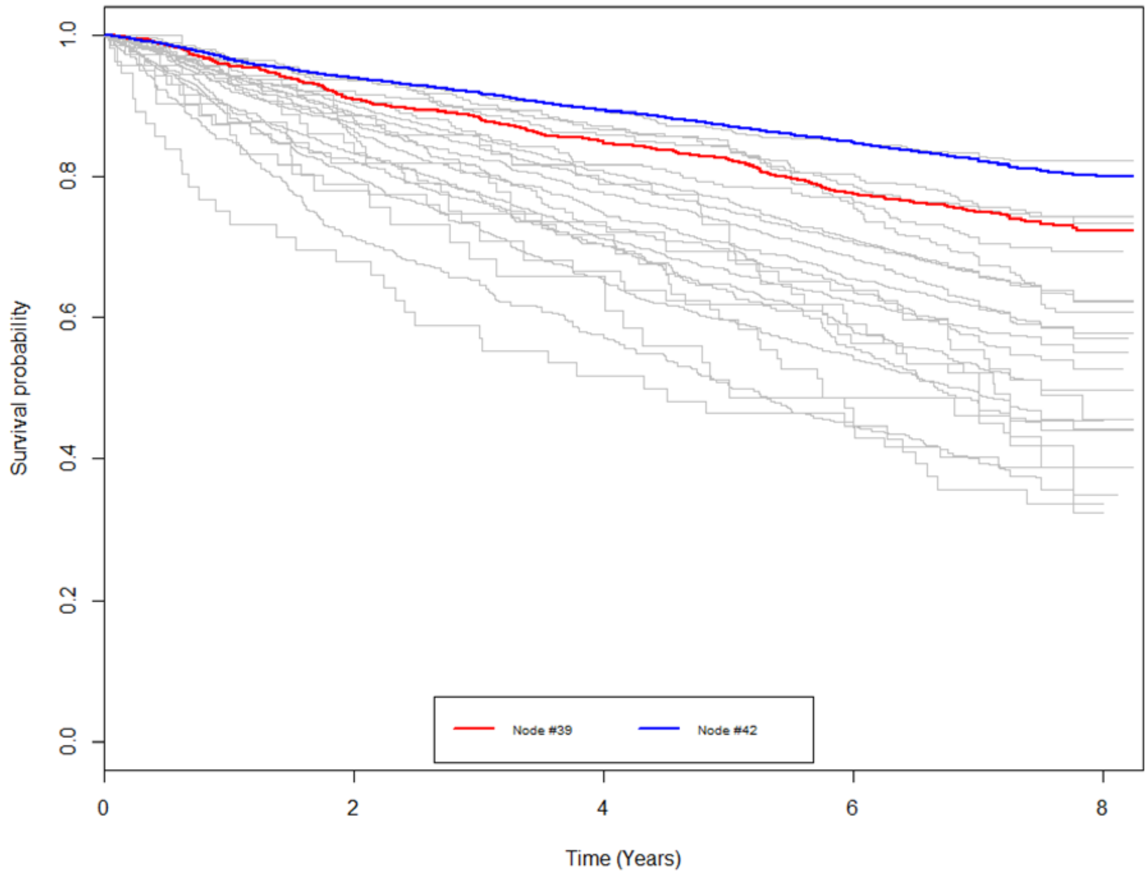



Figure 44. KM Survival Function Estimates for Females (Node 39), Males (Node 42), Age 44, and "Retirement" Military Separation Reason

# V. SUMMARY

This section presents a summary of our data understanding, preparation, and analysis. Recommendations for future work are also provided.

## A. CONCLUSIONS

### 1. Data Understanding

The PDE facilitates the analysis of sensitive DoD personnel data due to secure connectivity and the variety of analyst tools available. There is always room for improvement in terms of the quality of the data uploaded from DMDC into the PDE and its documentation, but the data used for the research was sufficient to meet the intent of gaining initial insights into DoD civilian attrition factors.

### 2. Data Preparation

By leveraging the master and transactions files that were available to us, we were able to construct a cohort of employees for the study. However, the 7% of employees who "disappear" during the study presented a significant challenge to fully understand why these particular employees departed.

### 3. Data Analysis

In comparison of blue-collar and white-collar employees, we see very similar trends among the survival probabilities of all the covariates analyzed between the two groups including salary, work level, military experience, annuitant status, branch of service, census bureau region, and race-ethnicity. We do, however, find that younger blue-collar males (29-years-old or less) have a higher survival probability than younger white-collar males. Blue-collar jobs are based more on experience than educational background, therefore younger blue-collar males may be more likely to stay employed with the DoD to increase their level of expertise within a particular trade. We also find blue-collar females, regardless of age, have a higher survival probability than white-collar females. This may be a result of the education level differences between the two groups. Blue-collar females may be more inclined to stay with the DoD due to a lack of quality jobs in the civilian sector that only require a high school diploma.

We also find that blue-collar and white-collar retired military members have the highest survival probability of the 2009 cohort and therefore have the potential for being good hires for the DoD. We do, however, find that employees with no military experience tend to have greater longevity than employees who served less than 20 years in the military, based on the survival probabilities. We are not advocating for the DoD to give stronger hiring preference to military retirees, but we do find it interesting that the presence of a military background does not necessarily translate to a lower attrition rate.

In addition, we find that 30% and 35% of the blue-collar and white-collar employees, respectively, attrite during the eight-year study period and the majority of employees separate within the first two years of employment. At the aggregate level, the probability of employee survivability increased among white-collar and blue-collar employees with families, higher salaries (greater than $50,000) and higher education (associate degree, bachelor's degree, or master's degree). White-collar professional and administrative employees and blue-collar and white-collar employees with mid-level or senior-level jobs also have a higher survival probability. Finally, white-collar and blue-collar employees who are male, who are between the ages of 35 and 54, or who work for the Navy have an increased survivability.

The findings from this research may be used by the OPA to better understand DoD civilian attrition factors, in order to implement DoD civilian employee policies that could reduce attrition. The study of attrition factors may also lead to improved models and tools to help better forecast DoD civilian attrition.

## B. FUTURE WORK

### 1. Exploration of Different Cohorts

The PDE also contains the data for multiple other cohorts of newly hired DoD employees beginning employment beyond 2009. These cohorts should also be examined to compare the findings from the 2009 cohort.

### 2. Differences in High- and Low-Risk Employees

The attrition factors of employees who resign versus employees who are terminated should be explored to determine which employees may be at higher risk for these two attrition causes. By exploring these differences, OPA may be able to construct policies to better screen DoD civilian employee applicants.

### 3. Examine Geography and Attrition

The examination of DoD civilian employees in more specific regions of the country, to include zip codes and military unit identification codes, should be conducted to compare which regions and military units are more at risk for attrition. We note that gaining access to unscrambled and unmasked zip codes and unit identification codes in the PDE requires an approved Initial Review Board protocol and several months of lead time.

THIS PAGE INTENTIONALLY LEFT BLANK

# APPENDIX A.  CENSUS BUREAU REGIONS AND DIVISIONS

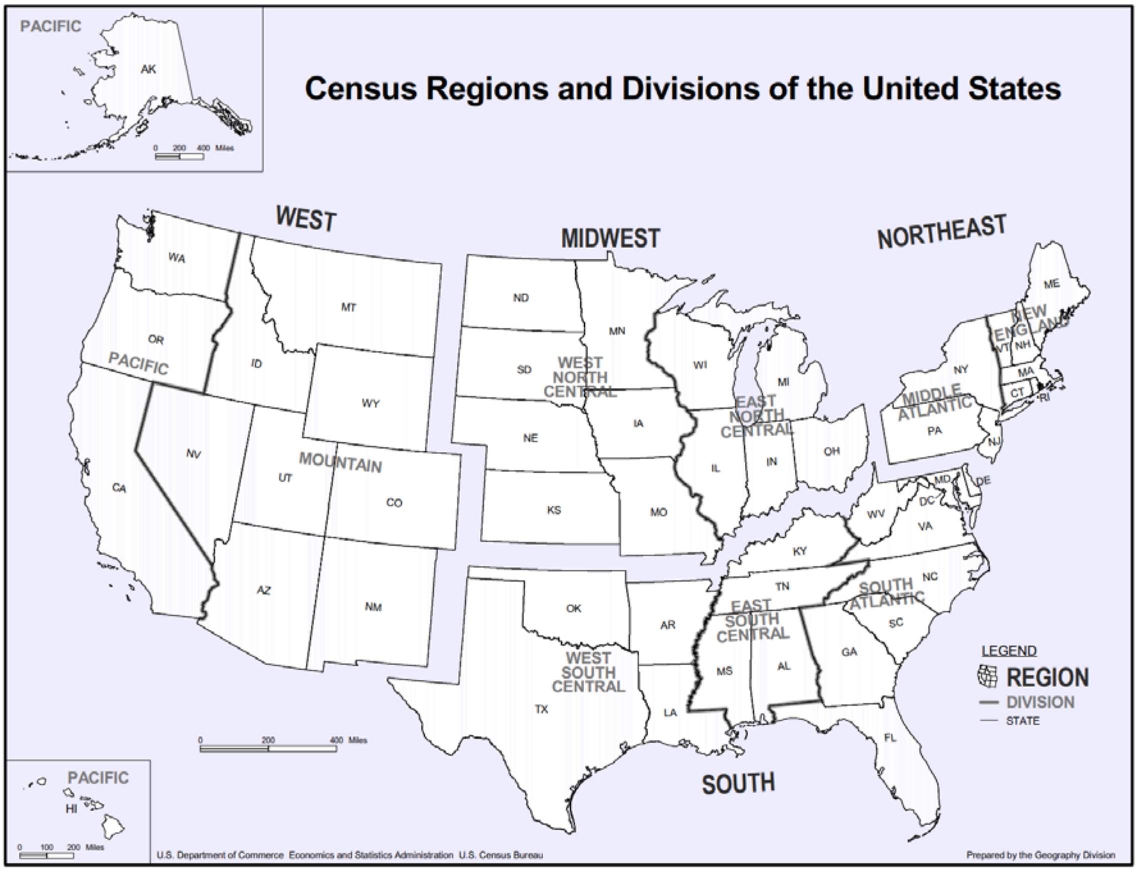



Figure A1. Census Bureau Regions and Division of the United States.
Source: Census Bureau (2010).

THIS PAGE INTENTIONALLY LEFT BLANK

# APPENDIX B. KAPLAN-MEIER ESTIMATED SURVIVAL FUNCTIONS AND SURVIVAL TREES

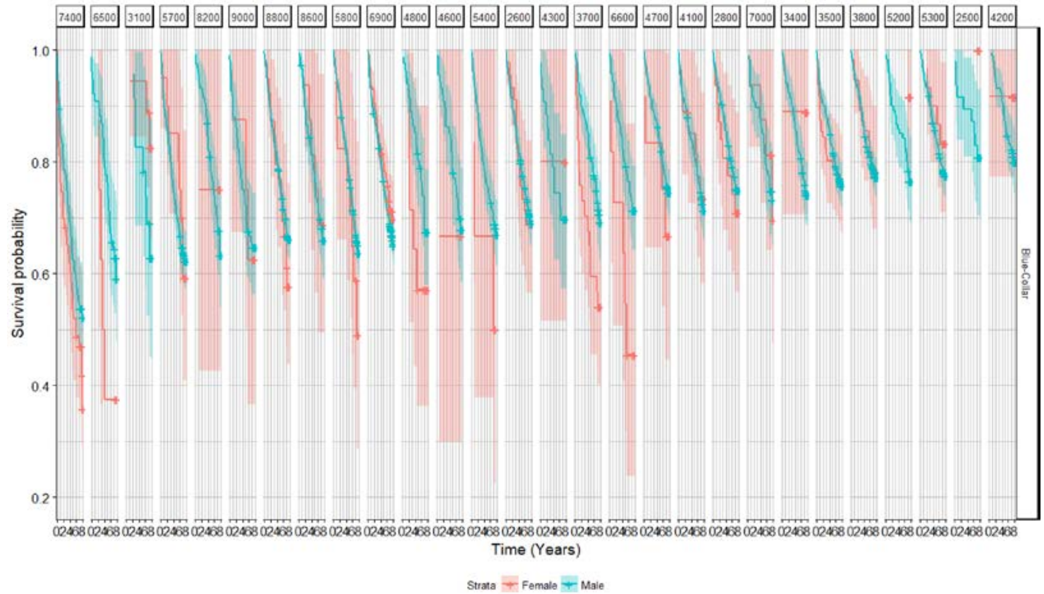



Figure B1. KM Estimates of Occupational Family by Gender and Collar Type

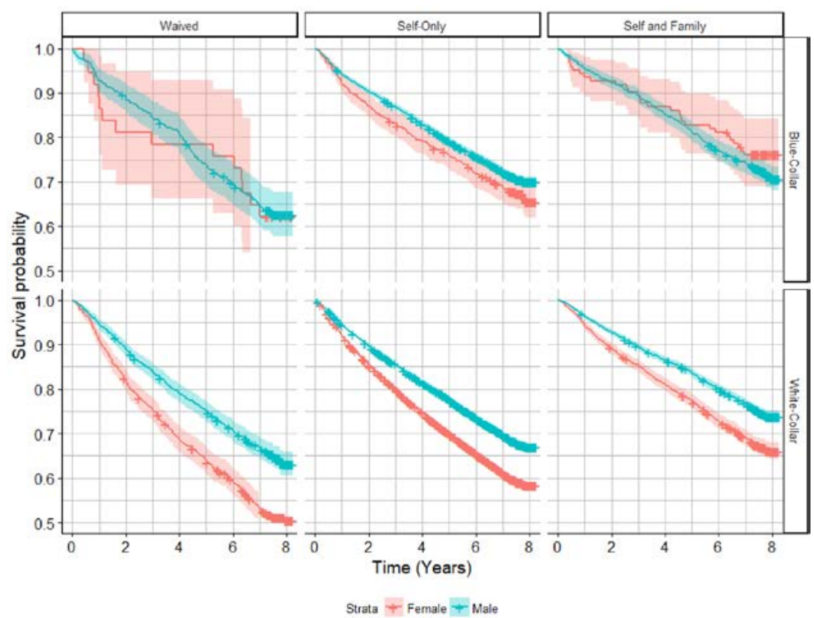



Figure B2. KM Estimates of Life Insurance Coverage by Gender and Collar Type

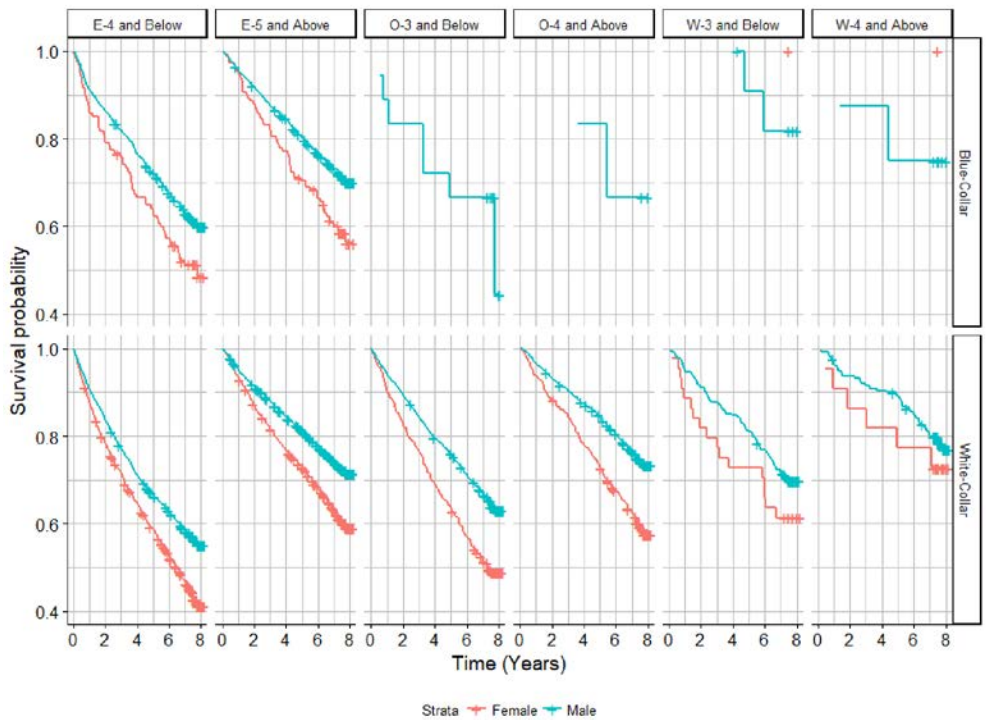



Figure B3. KM Estimates of Military Paygrades by Gender and Collar Type (without 95% Confidence Intervals)

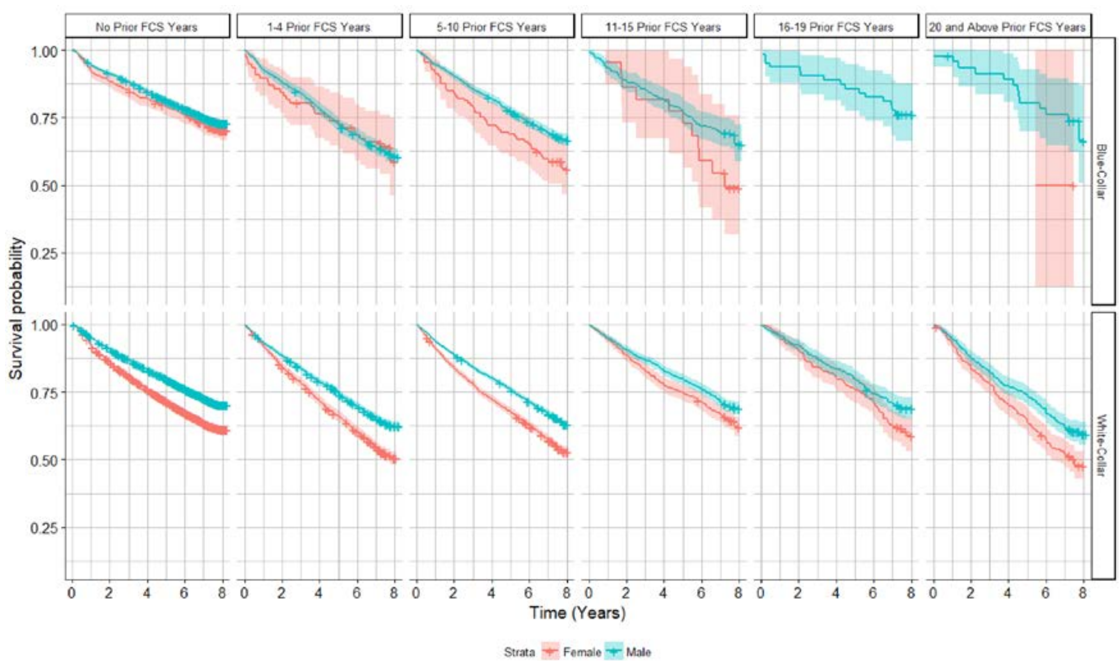



Figure B4. KM Estimates of Prior FCS Years by Gender and Collar Type

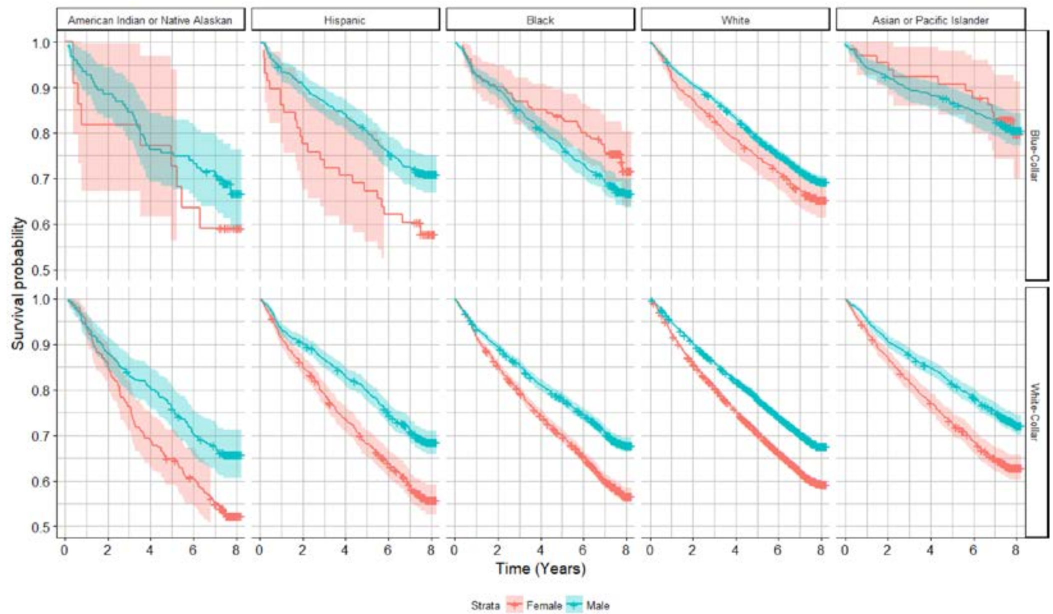



Figure B5. KM Estimates of Race-Ethnicity by Gender and Collar Type

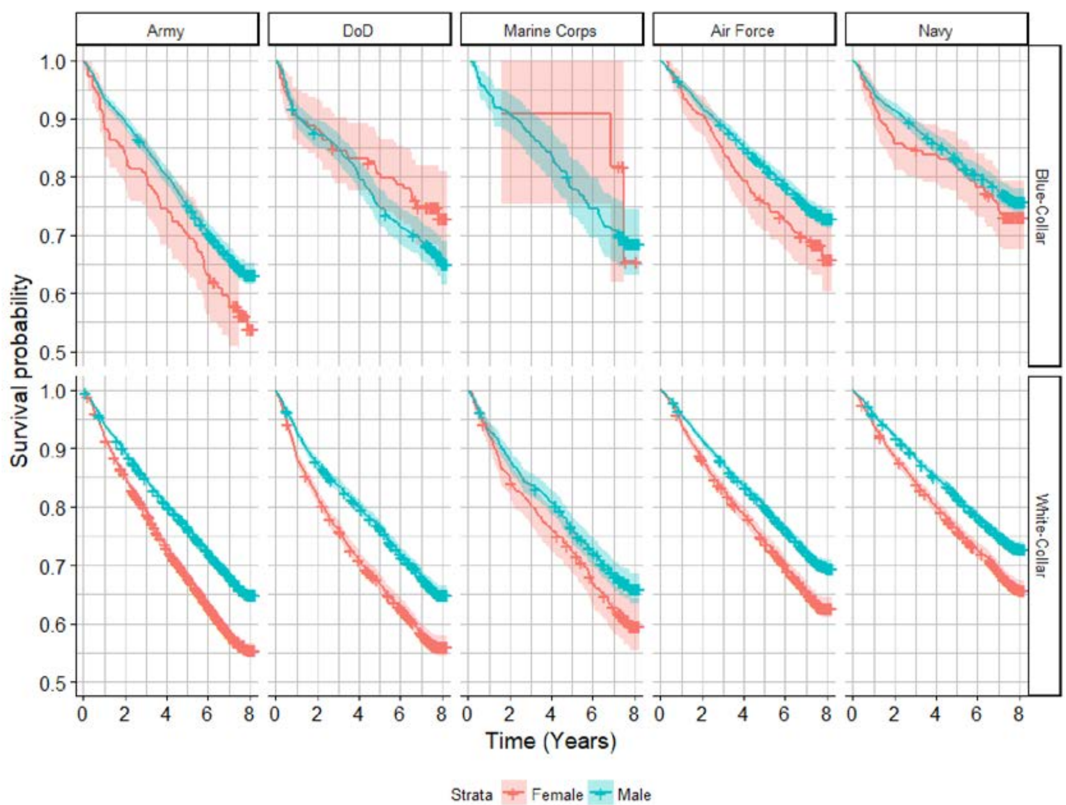



Figure B6. KM Estimates of Branch of Service by Gender and Collar Type

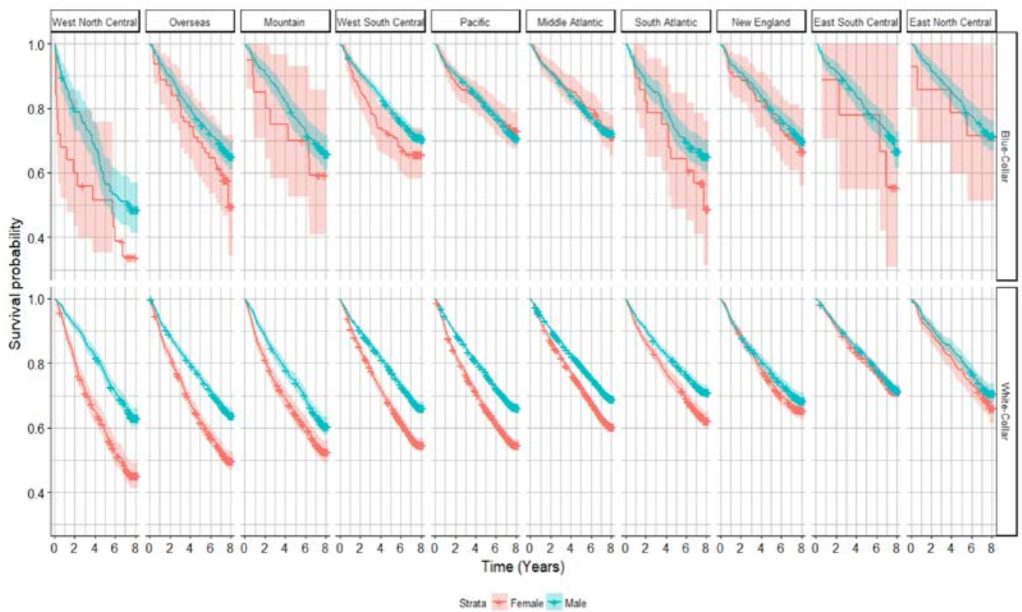



Figure B7 KM Estimates of Bureau of the Census Divisions by Gender and Collar Type

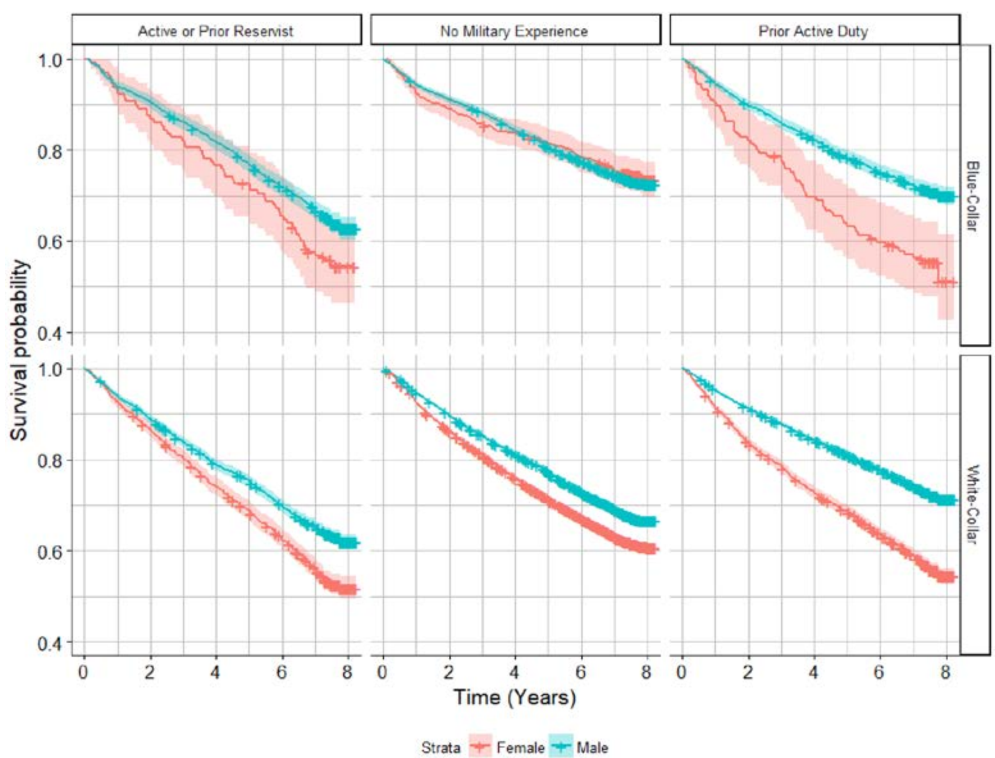



Figure B8. KM Estimates of Prior Military Experience by Gender and Collar Type

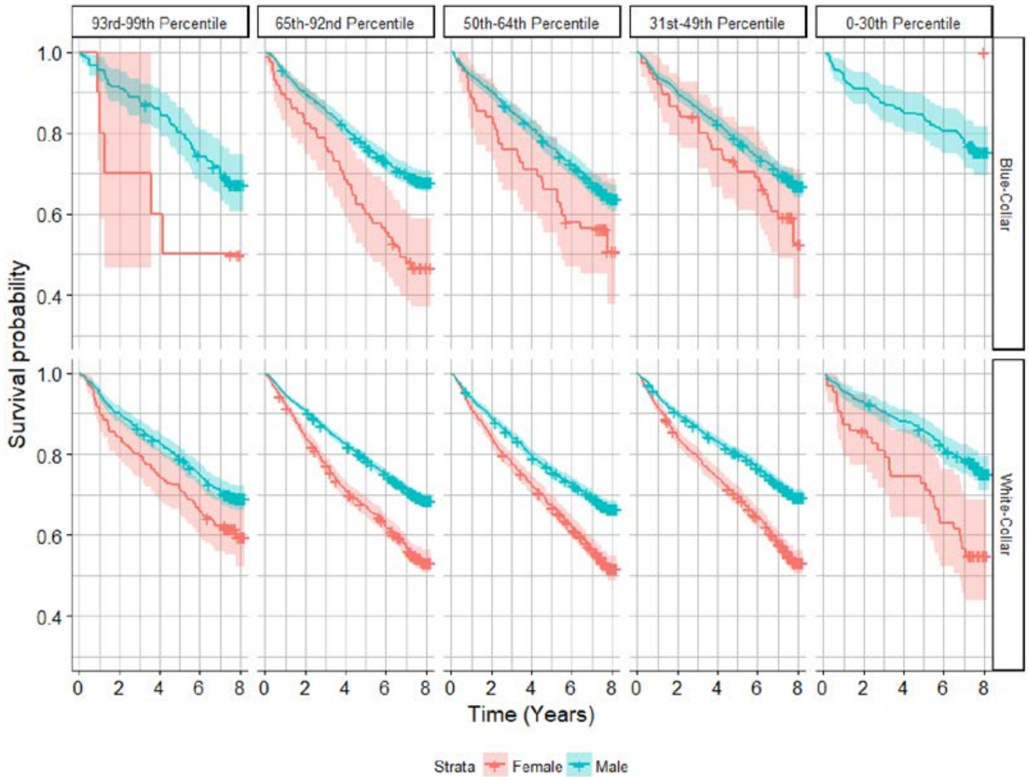



Figure B9. KM Estimates of AFQT Percentile Scores by Gender and Collar Type

Table B1. Full Survival Tree for 2009 Cohort

```
Model formula:
Surv(StartAge.new, EndAge.new, Event) ~ AGE + PRIOR_FCS_YEARS +
    GENDER + BRANCH_OF_SERVICE + ANNUITANT_STATUS + PRIOR_FCS_YEARS +
    MILITARY_EXPERIENCE + LIFE_INSURANCE + CENSUS_BUREAU_DIVISION +
    RACE_ETHNICITY + OCCUPATIONAL_CATEGORY + COLLAR_TYPE

Fitted party:
[1] root
|   [2] OCCUPATIONAL_CATEGORY in Clerical, Other White-Collar, Technical
|   |   [3] ANNUITANT_STATUS in Not Retired AD, RESERVIST, Civilian
|   |   |   [4] CENSUS_BUREAU_DIVISION in East North Central, East South Central, Middle Atlantic, New England, South Atlantic
|   |   |   |   [5] ANNUITANT_STATUS in Not Retired AD, RESERVIST
|   |   |   |   |   [6] OCCUPATIONAL_CATEGORY in Clerical, Other White-Collar: 7.253 (n = 1581)
|   |   |   |   |   [7] OCCUPATIONAL_CATEGORY in Technical
|   |   |   |   |   |   [8] BRANCH_OF_SERVICE in Army, Marine Corps, Air Force, Navy: Inf (n = 1730)
|   |   |   |   |   |   [9] BRANCH_OF_SERVICE in DoD: Inf (n = 162)
|   |   |   |   [10] ANNUITANT_STATUS in Civilian
|   |   |   |   |   [11] LIFE_INSURANCE in Waived, Self-Only
|   |   |   |   |   |   [12] OCCUPATIONAL_CATEGORY in Clerical, Other White-Collar
|   |   |   |   |   |   |   [13] BRANCH_OF_SERVICE in Army, Marine Corps, Air Force, Navy
|   |   |   |   |   |   |   |   [14] AGE <= 21: 5.095 (n = 500)
|   |   |   |   |   |   |   |   [15] AGE > 21
|   |   |   |   |   |   |   |   |   [16] PRIOR_FCS_YEARS <= 8: Inf (n = 1233)
|   |   |   |   |   |   |   |   |   [17] PRIOR_FCS_YEARS > 8: 7.253 (n = 73)
|   |   |   |   |   |   |   [18] BRANCH_OF_SERVICE in DoD
|   |   |   |   |   |   |   |   [19] CENSUS_BUREAU_DIVISION in East North Central: Inf (n = 217)
|   |   |   |   |   |   |   |   [20] CENSUS_BUREAU_DIVISION in East South Central, Middle Atlantic, New England, South Atlantic
|   |   |   |   |   |   |   |   |   [21] OCCUPATIONAL_CATEGORY in Clerical: 3.825 (n = 480)
|   |   |   |   |   |   |   |   |   [22] OCCUPATIONAL_CATEGORY in Other White-Collar: Inf (n = 238)
|   |   |   |   |   |   [23] OCCUPATIONAL_CATEGORY in Technical
|   |   |   |   |   |   |   [24] GENDER in Female: Inf (n = 1387)
|   |   |   |   |   |   |   [25] GENDER in Male: Inf (n = 674)
|   |   |   |   |   [26] LIFE_INSURANCE in Self-Plus-Family
|   |   |   |   |   |   [27] CENSUS_BUREAU_DIVISION in East North Central, East South Central: Inf (n = 257)
|   |   |   |   |   |   [28] CENSUS_BUREAU_DIVISION in Middle Atlantic, New England, South Atlantic: Inf (n = 466)
|   |   |   [29] CENSUS_BUREAU_DIVISION in Mountain, Overseas, Pacific, West North Central, West South Central
|   |   |   |   [30] BRANCH_OF_SERVICE in Army, DoD, Air Force
|   |   |   |   |   [31] BRANCH_OF_SERVICE in Army, Air Force
|   |   |   |   |   |   [32] AGE <= 26
|   |   |   |   |   |   |   [33] AGE <= 21: 3.529 (n = 267)
|   |   |   |   |   |   |   [34] AGE > 21
|   |   |   |   |   |   |   |   [35] OCCUPATIONAL_CATEGORY in Clerical, Technical
|   |   |   |   |   |   |   |   |   [36] CENSUS_BUREAU_DIVISION in Mountain, West North Central, West South Central: 6.428 (n = 472)
|   |   |   |   |   |   |   |   |   [37] CENSUS_BUREAU_DIVISION in Overseas, Pacific: 3.817 (n = 260)
|   |   |   |   |   |   |   |   [38] OCCUPATIONAL_CATEGORY in Other White-Collar: Inf (n = 393)
|   |   |   |   |   |   [39] AGE > 26
|   |   |   |   |   |   |   [40] RACE_ETHNICITY in American Indian or Native Alaskan, Hispanic, Asian or Pacific Islander: Inf (n = 645)
|   |   |   |   |   |   |   [41] RACE_ETHNICITY in Black, White: 7.754 (n = 3189)
|   |   |   |   |   [42] BRANCH_OF_SERVICE in DoD
|   |   |   |   |   |   [43] OCCUPATIONAL_CATEGORY in Clerical
|   |   |   |   |   |   |   [44] AGE <= 35
|   |   |   |   |   |   |   |   [45] GENDER in Female: 2.105 (n = 172)
|   |   |   |   |   |   |   |   [46] GENDER in Male: 3.135 (n = 95)
|   |   |   |   |   |   |   [47] AGE > 35: 5.232 (n = 231)
|   |   |   |   |   |   [48] OCCUPATIONAL_CATEGORY in Other White-Collar, Technical: Inf (n = 199)
|   |   |   |   [49] BRANCH_OF_SERVICE in Marine Corps, Navy
|   |   |   |   |   [50] AGE <= 30: Inf (n = 628)
|   |   |   |   |   [51] AGE > 30
|   |   |   |   |   |   [52] CENSUS_BUREAU_DIVISION in Mountain, Overseas: 6.760 (n = 59)
|   |   |   |   |   |   [53] CENSUS_BUREAU_DIVISION in Pacific, West North Central, West South Central: Inf (n = 681)
|   |   [54] ANNUITANT_STATUS in Retired Military Officer, Retired Enlisted Service Member
|   |   |   [55] OCCUPATIONAL_CATEGORY in Clerical, Other White-Collar
|   |   |   |   [56] MILITARY_EXPERIENCE in ACITVE DUTY: Inf (n = 788)
|   |   |   |   [57] MILITARY_EXPERIENCE in NONE, RESERVIST
|   |   |   |   |   [58] AGE <= 56: Inf (n = 150)
```

Table B1 (con’t.).    Full Survival Tree for 2009 Cohort

```
|   |   |   |   |   [59] AGE > 56: 3.250 (n = 60)
|   |   |   [60] OCCUPATIONAL_CATEGORY in Technical
|   |   |   |   [61] MILITARY_EXPERIENCE in ACITVE DUTY
|   |   |   |   |   [62] GENDER in Female: Inf (n = 324)
|   |   |   |   |   [63] GENDER in Male
|   |   |   |   |   |   [64] CENSUS_BUREAU_DIVISION in East North Central, East South Central, Middle Atlantic, Mountain, New England, Pacific, South Atlantic: Inf (n =
1017)
|   |   |   |   |   |   [65] CENSUS_BUREAU_DIVISION in Overseas, West North Central, West South Central: Inf (n = 283)
|   |   |   |   [66] MILITARY_EXPERIENCE in NONE, RESERVIST: Inf (n = 436)
|   [67] OCCUPATIONAL_CATEGORY in Professional, Administrative, Blue-Collar
|   |   [68] ANNUITANT_STATUS in Not Retired AD, RESERVIST, Civilian, Retired Military Officer
|   |   |   [69] BRANCH_OF_SERVICE in Army, DoD
|   |   |   |   [70] CENSUS_BUREAU_DIVISION in East North Central, East South Central, Middle Atlantic
|   |   |   |   |   [71] AGE <= 58
|   |   |   |   |   |   [72] ANNUITANT_STATUS in Not Retired AD, RESERVIST
|   |   |   |   |   |   |   [73] PRIOR_FCS_YEARS <= 4: Inf (n = 624)
|   |   |   |   |   |   |   [74] PRIOR_FCS_YEARS > 4
|   |   |   |   |   |   |   |   [75] GENDER in Female: Inf (n = 114)
|   |   |   |   |   |   |   |   [76] GENDER in Male: Inf (n = 600)
|   |   |   |   |   |   [77] ANNUITANT_STATUS in Civilian, Retired Military Officer
|   |   |   |   |   |   |   [78] PRIOR_FCS_YEARS <= 19: Inf (n = 3969)
|   |   |   |   |   |   |   [79] PRIOR_FCS_YEARS > 19: 7.253 (n = 57)
|   |   |   |   |   [80] AGE > 58: 5.509 (n = 182)
|   |   |   |   [81] CENSUS_BUREAU_DIVISION in Mountain, New England, Overseas, Pacific, South Atlantic, West North Central, West South Central
|   |   |   |   |   [82] GENDER in Female
|   |   |   |   |   |   [83] OCCUPATIONAL_CATEGORY in Administrative, Blue-Collar
|   |   |   |   |   |   |   [84] PRIOR_FCS_YEARS <= 15
|   |   |   |   |   |   |   |   [85] ANNUITANT_STATUS in Not Retired AD, RESERVIST: Inf (n = 553)
|   |   |   |   |   |   |   |   [86] ANNUITANT_STATUS in Civilian, Retired Military Officer
|   |   |   |   |   |   |   |   |   [87] LIFE_INSURANCE in Waived: 6.546 (n = 48)
|   |   |   |   |   |   |   |   |   [88] LIFE_INSURANCE in Self-Only, Self-Plus-Family
|   |   |   |   |   |   |   |   |   |   [89] CENSUS_BUREAU_DIVISION in Overseas: 5.974 (n = 51)
|   |   |   |   |   |   |   |   |   |   [90] CENSUS_BUREAU_DIVISION in Mountain, New England, Pacific, South Atlantic, West North Central, West South Central: Inf (n =
1072)
|   |   |   |   |   |   |   [91] PRIOR_FCS_YEARS > 15: 6.752 (n = 184)
|   |   |   |   |   |   [92] OCCUPATIONAL_CATEGORY in Professional
|   |   |   |   |   |   |   [93] CENSUS_BUREAU_DIVISION in New England, Pacific, South Atlantic, West North Central, West South Central
|   |   |   |   |   |   |   |   [94] AGE <= 57
|   |   |   |   |   |   |   |   |   [95] MILITARY_EXPERIENCE in ACITVE DUTY, RESERVIST
|   |   |   |   |   |   |   |   |   |   [96] AGE <= 31
|   |   |   |   |   |   |   |   |   |   |   [97] PRIOR_FCS_YEARS <= 3: 6.341 (n = 31)
|   |   |   |   |   |   |   |   |   |   |   [98] PRIOR_FCS_YEARS > 3: 2.782 (n = 45)
|   |   |   |   |   |   |   |   |   |   [99] AGE > 31: Inf (n = 246)
|   |   |   |   |   |   |   |   |   [100] MILITARY_EXPERIENCE in NONE: Inf (n = 2038)
|   |   |   |   |   |   |   |   [101] AGE > 57: 4.323 (n = 99)
|   |   |   |   |   |   |   [102] CENSUS_BUREAU_DIVISION in Mountain, Overseas
|   |   |   |   |   |   |   |   [103] BRANCH_OF_SERVICE in Army: 4.290 (n = 261)
|   |   |   |   |   |   |   |   [104] BRANCH_OF_SERVICE in DoD: Inf (n = 139)
|   |   |   |   |   [105] GENDER in Male
|   |   |   |   |   |   [106] PRIOR_FCS_YEARS <= 23
|   |   |   |   |   |   |   [107] LIFE_INSURANCE in Waived, Self-Only
|   |   |   |   |   |   |   |   [108] MILITARY_EXPERIENCE in ACITVE DUTY
|   |   |   |   |   |   |   |   |   [109] ANNUITANT_STATUS in Not Retired AD
|   |   |   |   |   |   |   |   |   |   [110] COLLAR_TYPE in Blue-Collar: Inf (n = 345)
|   |   |   |   |   |   |   |   |   |   [111] COLLAR_TYPE in White-Collar: Inf (n = 946)
|   |   |   |   |   |   |   |   |   [112] ANNUITANT_STATUS in Retired Military Officer
|   |   |   |   |   |   |   |   |   |   [113] AGE <= 54: Inf (n = 609)
|   |   |   |   |   |   |   |   |   |   [114] AGE > 54: Inf (n = 86)
|   |   |   |   |   |   |   |   [115] MILITARY_EXPERIENCE in NONE, RESERVIST
|   |   |   |   |   |   |   |   |   [116] AGE <= 55
|   |   |   |   |   |   |   |   |   |   [117] LIFE_INSURANCE in Waived: Inf (n = 271)
|   |   |   |   |   |   |   |   |   |   [118] LIFE_INSURANCE in Self-Only
|   |   |   |   |   |   |   |   |   |   |   [119] CENSUS_BUREAU_DIVISION in Mountain, Overseas: Inf (n = 533)
|   |   |   |   |   |   |   |   |   |   |   [120] CENSUS_BUREAU_DIVISION in New England, Pacific, South Atlantic, West North Central, West South Central: Inf (n = 3711)
|   |   |   |   |   |   |   |   |   [121] AGE > 55
|   |   |   |   |   |   |   |   |   |   [122] AGE <= 60
|   |   |   |   |   |   |   |   |   |   |   [123] COLLAR_TYPE in Blue-Collar: 4.999 (n = 56)
|   |   |   |   |   |   |   |   |   |   |   [124] COLLAR_TYPE in White-Collar: Inf (n = 304)
|   |   |   |   |   |   |   |   |   |   [125] AGE > 60: 5.196 (n = 205)
|   |   |   |   |   |   |   [126] LIFE_INSURANCE in Self-Plus-Family: Inf (n = 1553)
|   |   |   |   |   |   [127] PRIOR_FCS_YEARS > 23
|   |   |   |   |   |   |   [128] AGE <= 51: 7.754 (n = 38)
|   |   |   |   |   |   |   [129] AGE > 51: 3.751 (n = 75)
|   |   |   [130] BRANCH_OF_SERVICE in Marine Corps, Air Force, Navy
|   |   |   |   [131] ANNUITANT_STATUS in Not Retired AD, RESERVIST
|   |   |   |   |   [132] AGE <= 35
|   |   |   |   |   |   [133] GENDER in Female: Inf (n = 637)
|   |   |   |   |   |   [134] GENDER in Male
|   |   |   |   |   |   |   [135] OCCUPATIONAL_CATEGORY in Administrative: Inf (n = 924)
|   |   |   |   |   |   |   [136] OCCUPATIONAL_CATEGORY in Professional, Blue-Collar: Inf (n = 1762)
```

## Table B1 (con't.). Full Survival Tree for 2009 Cohort

```
|   |   |   |   |   [137] AGE > 35
|   |   |   |   |   |   [138] GENDER in Female: Inf (n = 341)
|   |   |   |   |   |   [139] GENDER in Male
|   |   |   |   |   |   |   [140] AGE <= 57: Inf (n = 1627)
|   |   |   |   |   |   |   [141] AGE > 57: 6.335 (n = 28)
|   |   |   |   [142] ANNUITANT_STATUS in Civilian, Retired Military Officer
|   |   |   |   |   [143] AGE <= 55
|   |   |   |   |   |   [144] LIFE_INSURANCE in Waived, Self-Only
|   |   |   |   |   |   |   [145] OCCUPATIONAL_CATEGORY in Administrative, Blue-Collar
|   |   |   |   |   |   |   |   [146] CENSUS_BUREAU_DIVISION in East North Central, Middle Atlantic, New England, Pacific, South Atlantic, West North Central, West
South Central: Inf (n = 4777)
|   |   |   |   |   |   |   |   [147] CENSUS_BUREAU_DIVISION in East South Central, Mountain, Overseas: Inf (n = 374)
|   |   |   |   |   |   |   [148] OCCUPATIONAL_CATEGORY in Professional
|   |   |   |   |   |   |   |   [149] GENDER in Female
|   |   |   |   |   |   |   |   |   [150] CENSUS_BUREAU_DIVISION in East North Central, East South Central, Middle Atlantic, New England, West North Central: Inf (n =
316)
|   |   |   |   |   |   |   |   |   [151] CENSUS_BUREAU_DIVISION in Mountain, Overseas, Pacific, South Atlantic, West South Central: Inf (n = 1292)
|   |   |   |   |   |   |   |   [152] GENDER in Male
|   |   |   |   |   |   |   |   |   [153] RACE_ETHNICITY in American Indian or Native Alaskan, Hispanic, Asian or Pacific Islander
|   |   |   |   |   |   |   |   |   |   [154] BRANCH_OF_SERVICE in Navy: Inf (n = 362)
|   |   |   |   |   |   |   |   |   |   [155] BRANCH_OF_SERVICE in Marine Corps, Air Force: Inf (n = 129)
|   |   |   |   |   |   |   |   |   [156] RACE_ETHNICITY in Black, White
|   |   |   |   |   |   |   |   |   |   [157] CENSUS_BUREAU_DIVISION in East North Central, East South Central, Middle Atlantic, Mountain, New England, Overseas, South
Atlantic, West North Central: Inf (n = 2377)
|   |   |   |   |   |   |   |   |   |   [158] CENSUS_BUREAU_DIVISION in Pacific, West South Central: Inf (n = 639)
|   |   |   |   |   |   [159] LIFE_INSURANCE in Self-Plus-Family: Inf (n = 1753)
|   |   |   |   |   [160] AGE > 55
|   |   |   |   |   |   [161] AGE <= 58: Inf (n = 438)
|   |   |   |   |   |   [162] AGE > 58: 5.944 (n = 441)
|   |   [163] ANNUITANT_STATUS in Retired Enlisted Service Member
|   |   |   [164] MILITARY_EXPERIENCE in ACITVE DUTY, RESERVIST
|   |   |   |   [165] CENSUS_BUREAU_DIVISION in East North Central, East South Central, New England, South Atlantic, West South Central: Inf (n = 3491)
|   |   |   |   [166] CENSUS_BUREAU_DIVISION in Middle Atlantic, Mountain, Overseas, Pacific, West North Central: Inf (n = 1721)
|   |   |   [167] MILITARY_EXPERIENCE in NONE
|   |   |   |   [168] AGE <= 57
|   |   |   |   |   [169] COLLAR_TYPE in Blue-Collar: Inf (n = 185)
|   |   |   |   |   [170] COLLAR_TYPE in White-Collar: Inf (n = 596)
|   |   |   |   [171] AGE > 57: 6.001 (n = 155)

Number of inner nodes:    85
Number of terminal nodes: 86
```

## Table B2. Pruned Survival Tree for 2009 Cohort

```
Model formula:
Surv(StartAge.new, EndAge.new, Event) ~ AGE + PRIOR_FCS_YEARS +
    GENDER + BRANCH_OF_SERVICE + ANNUITANT_STATUS + PRIOR_FCS_YEARS +
    MILITARY_EXPERIENCE + LIFE_INSURANCE + CENSUS_BUREAU_DIVISION +
    RACE_ETHNICITY + OCCUPATIONAL_CATEGORY + COLLAR_TYPE

Fitted party:
[1] root
|   [2] OCCUPATIONAL_CATEGORY in Clerical, Other White-Collar, Technical
|   |   [3] ANNUITANT_STATUS in Not Retired AD, RESERVIST, Civilian
|   |   |   [4] CENSUS_BUREAU_DIVISION in East North Central, East South Central, Middle Atlantic, New England, South Atlantic: Inf (n = 8998)
|   |   |   [5] CENSUS_BUREAU_DIVISION in Mountain, Overseas, Pacific, West North Central, West South Central
|   |   |   |   [6] BRANCH_OF_SERVICE in Army, DoD, Air Force: 7.006 (n = 5923)
|   |   |   |   [7] BRANCH_OF_SERVICE in Marine Corps, Navy: Inf (n = 1368)
|   |   [8] ANNUITANT_STATUS in Retired Military Officer, Retired Enlisted Service Member: Inf (n = 3058)
|   [9] OCCUPATIONAL_CATEGORY in Professional, Administrative, Blue-Collar
|   |   [10] ANNUITANT_STATUS in Not Retired AD, RESERVIST, Civilian, Retired Military Officer
|   |   |   [11] BRANCH_OF_SERVICE in Army, DoD
|   |   |   |   [12] CENSUS_BUREAU_DIVISION in East North Central, East South Central, Middle Atlantic: Inf (n = 5546)
|   |   |   |   [13] CENSUS_BUREAU_DIVISION in Mountain, New England, Overseas, Pacific, South Atlantic, West North Central, West South Central
|   |   |   |   |   [14] GENDER in Female: Inf (n = 4767)
|   |   |   |   |   [15] GENDER in Male: Inf (n = 8732)
|   |   |   [16] BRANCH_OF_SERVICE in Marine Corps, Air Force, Navy
|   |   |   |   [17] ANNUITANT_STATUS in Not Retired AD, RESERVIST: Inf (n = 5319)
|   |   |   |   [18] ANNUITANT_STATUS in Civilian, Retired Military Officer: Inf (n = 12898)
|   |   [19] ANNUITANT_STATUS in Retired Enlisted Service Member
|   |   |   [20] MILITARY_EXPERIENCE in ACITVE DUTY, RESERVIST: Inf (n = 5212)
|   |   |   [21] MILITARY_EXPERIENCE in NONE: Inf (n = 936)

Number of inner nodes:    10
Number of terminal nodes: 11
```

Table B3. Full Survival Tree for Active Component and Reserve Component DoD Employees

```
Model formula:
Surv(StartAge.new, EndAge.new, Event) ~ AGE + GENDER + ANNUITANT_STATUS +
    MILITARY_EXPERIENCE + AFQT_PERCENTILE + MILITARY_PAYGRADE_GROUP +
    SEPARATION_REASON

Fitted party:
[1] root
|   [2] SEPARATION_REASON in Bad Behavior / Poor Performance, Expiration of Service, Other, NA, Active Reservist
|   |   [3] GENDER in Female
|   |   |   [4] AGE <= 31
|   |   |   |   [5] AGE <= 25
|   |   |   |   |   [6] MILITARY_EXPERIENCE in Prior Active Duty: 5.027 (n = 446)
|   |   |   |   |   [7] MILITARY_EXPERIENCE in Prior or Active Reservist: 6.669 (n = 298)
|   |   |   |   [8] AGE > 25
|   |   |   |   |   [9] SEPARATION_REASON in Bad Behavior / Poor Performance, Expiration of Service, Other: 6.869 (n = 1332)
|   |   |   |   |   [10] SEPARATION_REASON in NA, Active Reservist: Inf (n = 420)
|   |   |   [11] AGE > 31
|   |   |   |   [12] MILITARY_PAYGRADE_GROUP in EJ, OJ, WS: 7.504 (n = 739)
|   |   |   |   [13] MILITARY_PAYGRADE_GROUP in ES, OS, WJ: Inf (n = 1619)
|   |   [14] GENDER in Male
|   |   |   [15] MILITARY_PAYGRADE_GROUP in EJ
|   |   |   |   [16] MILITARY_EXPERIENCE in Prior Active Duty
|   |   |   |   |   [17] ANNUITANT_STATUS in No Annuitant Status: Inf (n = 1945)
|   |   |   |   |   [18] ANNUITANT_STATUS in Retired Enlisted Service Member: 4.323 (n = 56)
|   |   |   |   [19] MILITARY_EXPERIENCE in Prior or Active Reservist
|   |   |   |   |   [20] SEPARATION_REASON in Bad Behavior / Poor Performance, NA, Active Reservist: Inf (n = 1202)
|   |   |   |   |   [21] SEPARATION_REASON in Expiration of Service, Other: Inf (n = 318)
|   |   |   [22] MILITARY_PAYGRADE_GROUP in ES, OJ, OS, WJ, WS
|   |   |   |   [23] ANNUITANT_STATUS in No Annuitant Status
|   |   |   |   |   [24] SEPARATION_REASON in Bad Behavior / Poor Performance, Expiration of Service, Other
|   |   |   |   |   |   [25] SEPARATION_REASON in Bad Behavior / Poor Performance: 7.124 (n = 71)
|   |   |   |   |   |   [26] SEPARATION_REASON in Expiration of Service, Other: Inf (n = 5064)
|   |   |   |   |   [27] SEPARATION_REASON in NA, Active Reservist: Inf (n = 3243)
|   |   |   |   [28] ANNUITANT_STATUS in Retired Military Officer, Retired Enlisted Service Member
|   |   |   |   |   [29] MILITARY_EXPERIENCE in Prior Active Duty
|   |   |   |   |   |   [30] SEPARATION_REASON in NA, Active Reservist: Inf (n = 648)
|   |   |   |   |   |   [31] SEPARATION_REASON in Expiration of Service, Other
|   |   |   |   |   |   |   [32] AGE <= 30: 4.999 (n = 41)
|   |   |   |   |   |   |   [33] AGE > 30: Inf (n = 341)
|   |   |   |   |   [34] MILITARY_EXPERIENCE in Prior or Active Reservist
|   |   |   |   |   |   [35] AGE <= 54: Inf (n = 368)
|   |   |   |   |   |   [36] AGE > 54: 5.749 (n = 72)
|   [37] SEPARATION_REASON in Retirement
|   |   [38] GENDER in Female
|   |   |   [39] AGE <= 53: Inf (n = 1170)
|   |   |   [40] AGE > 53: 7.504 (n = 72)
|   |   [41] GENDER in Male
|   |   |   [42] AGE <= 53: Inf (n = 7832)
|   |   |   [43] AGE > 53
|   |   |   |   [44] AGE <= 58: Inf (n = 342)
|   |   |   |   [45] AGE > 58: 7.001 (n = 71)

Number of inner nodes:    22
Number of terminal nodes: 23
```

THIS PAGE INTENTIONALLY LEFT BLANK

# APPENDIX C. TABLES OF OCCUPATIONAL GROUPS AND FAMILIES

Table C1. White-Collar Occupational Groups Numbers and Names

| Occupational Group Number | Occupational Group Name |
|---|---|
| 0000 | Miscellaneous Occupations |
| 0100 | Social, Science, Psychology, and Welfare |
| 0200 | Human Resources Management |
| 0300 | Administration, Operations, and General Management |
| 0400 | Natural Resources Management and Biological Sciences |
| 0500 | Accounting and Budget |
| 0600 | Medical, Hospital, Dental, and Public Health |
| 0800 | Engineering and Architecture |
| 0900 | Legal and Kindred |
| 1000 | Information and Arts |
| 1100 | Business and Industry |
| 1300 | Physical Sciences |
| 1400 | Library and Archives |
| 1500 | Mathematical Sciences |
| 1600 | Equipment, Facilities, and Services |
| 1700 | Education |
| 1800 | Inspection, Investigation, Enforcement, and Compliance |
| 1900 | Quality Assurance, Inspection, and Grading |
| 2000 | Supply |
| 2100 | Transportation |
| 2200 | Information Technology |

Table C2. Blue-Collar Occupational Families Numbers and Names

| Occupational Family Number | Occupational Family Name |
|---|---|
| 2500 | Wire Communications Equipment Installation and Equipment |
| 2600 | Electronic Equipment Installation And Maintenance |
| 2800 | Electrical Installation and Maintenance |
| 3100 | Fabric And Leather Work |
| 3400 | Machine Tool Work |
| 3500 | General Services and Support Work |
| 3700 | Metal Processing |
| 3800 | Metal Work |
| 4100 | Painting And Paperhanging |
| 4200 | Plumbing And Pipefitting |
| 4300 | Pliable Materials Work |
| 4600 | Wood Work |
| 4700 | General Maintenance And Operations Work |
| 4800 | General Equipment Maintenance |
| 5200 | Miscellaneous Occupations |
| 5300 | Industrial Equipment Maintenance |
| 5400 | Industrial Equipment Operation |
| 5700 | Transportation/Mobile Equipment Operation |
| 5800 | Transportation/Mobile Equipment Maintenance |
| 6500 | Ammunition, Explosives, and Toxic Materials |
| 6600 | Armament Work |
| 6900 | Warehousing And Stock Handling |
| 7000 | Packing And Processing |
| 7400 | Food Preparation And Serving |
| 8200 | Fluid Systems Maintenance |
| 8600 | Engine Overhaul |
| 8800 | Aircraft Overhaul |
| 9000 | Film Processing |

# LIST OF REFERENCES


Asch BJ, Mattock MG, Hosek J (2014) The federal civil service workforce assessing the effects on retention of pay freezes, unpaid furloughs, and other federal employee compensation changes in the Department of Defense. Report RR514, RAND Corporation, Santa Monica, CA. https://www.rand.org/pubs/research_reports/RR514.html.

Breiman L, Friedman J, Olshen RA, Stone CJ (1984) *Classification and Regression Trees* (Wadsworth International Group, Belmont, CA).

Buttrey SE, Clark D (2017) Civilian survival analysis of Navy Manpower Data. Unpublished technical report.

Buttrey SE, Klingensmith K, Whitaker LR (2018) Civilian Department of Defense attrition. Unpublished technical report.

Census Bureau (2010) Census regions and divisions of the United States. Accessed February 12, 2019, https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf.

Defense Manpower Data Center (2010), Memorandum of Agreement between Army Data Center Fairfield and Defense Manpower Data Center, July 2010.

Equal Employment Opportunity Commission (2009) FY 2009 Annual report on the federal workforce. Accessed February 3, 2019, https://www.eeoc.gov/federal/reports/fsp2009/index.cfm.

Fu W, Simonoff J (2017) LTRCtrees: Survival trees to fit left-truncated and right-censored and interval-censored survival data. R package version 0.5.0 Accessed February 12, 2019, https://CRAN.R-project.org/package=LTRCtrees.

Goodman CJ, Mance SM (2011) Employment loss and the 2007–09 recession: an overview. Accessed March 12, 2019 https://www.bls.gov/opub/mlr/2011/04/art1full.pdf.

Government Accountability Office (1997) Military attrition: DoD could save millions by better screening enlisted personnel. Accessed January 15, 2019, https://www.gao.gov/assets/160/155698.pdf.

Government Accountability Office (1998) Military attrition: “DoD” needs to better analyze reasons for separation and improve recruiting systems. Accessed January 15, 2019, https://www.gao.gov/assets/110/107274.pdf.

Hothorn T, Zeileis A (2015) partykit: A modular toolkit for recursive portioning in R. *Journal of Machine Learning Research*, 16: 3905–3909.

Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association.* 53(282): 457–481.

Kassambara A, Kosinski M (2017) survminer: Drawing Survival Curves using ‘gglpot2’. R package version 0.4.1. Accessed January 15, 2019, https://CRAN.R-project.org/package-survminer.

Kleinbaum DG, Klein M (2005) *Survival Analysis: A Self-Learning Text*, 2nd ed. (Springer, New York, NY).

Knapik JJ, Jones BH, Hauret K, Darakjy S, Piskator E (2004) A review of the literature on attrition from the military services: Risk factors for attrition and strategies to reduce attrition. Accessed February 1, 2019, https://apps.dtic.mil/dtic/tr/fulltext/u2/a427744.pdf.

Knapp D, Asch BJ, DeMartini C, Ruder T, Hanley J (2018) Using the Person-Event Data Environment for military personnel research in the Department of Defense: An evaluation of capability and potential uses. Report RR2302, RAND Corporation, Santa Monica, CA. https://www.rand.org/pubs/research_reports/RR2302.html.

Knapp D, Asch BJ, Mattock MG, Hosek J (2016) An enhanced capability to model how compensation policy affects U.S. Department of Defense civil service retention and cost. Report RR1503, RAND Corporation, Santa Monica, CA. https://www.rand.org/pubs/research_reports/RR1503.html.

Light P (2011) The campaign for high performance government, creating a high-performance government: A once-in-a-generation opportunity, New York University, Robert F. Wagner School of Public Service, June 21, 2011.

Martin T (1995) Who stays, who leaves? An analysis of first-term Army attrition. Doctoral dissertation, RAND Graduate School, California. https://www.rand.org/content/dam/rand/pubs/rgs_dissertations/2006/RGSD114.pdf.

Novak JD, Cañas AJ (2008) The Theory underlying concept maps and how to construct and use them. Accessed January 15, 2019, http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf.

Office of People Analytics (2017) Welcome to the Office of People Analytics. Accessed January 13, 2019, https://opa.defense.gov/.

Office of Personnel Management (1999) Appendix F. Accessed February 1, 2019, https://www.opm.gov/policy-data-oversight/pay-leave/pay-systems/federal-wage-system/appropriated-fund-operating-manual/appendixf.pdf.

Office of Personnel Management (2000) Subchapter S3. Accessed February 1, 2019, https://www.opm.gov/policy-data-oversight/pay-leave/pay-systems/federal-wage-system/nonappropriated-fund-operating-manual/subchapter3.pdf.

Office of Personnel Management (2002) Cost of administering the federal wage system. Accessed February 1, 2019, https://www.opm.gov/policy-data-oversight/pay-leave/pay-systems/federal-wage-system/reports-to-congress/cost-of-administering-the-federal-wage-system/.

Office of Personnel Management (2009) Introduction to the position classification standards. Accessed February 1, 2019, https://www.opm.gov/policy-data-oversight/classification-qualifications/classifying-general-schedule-positions/positionclassificationintro.pdf.

Office of Personnel Management (2010) Federal employment retirement system information: Eligibility. Accessed February 1, 2019, https://www.opm.gov/retirement-services/fers-information/eligibility/.

Office of Personnel Management (2013) Profile of federal civilian non-seasonal full-time employees. Accessed (February 7, 2019), https://www.opm.gov/policy-data-oversight/data-analysis-documentation/federal-employment-reports/reports-publications/profile-of-federal-civilian-non-postal-employees/.

Office of Personnel Management (2014) Chapter 31. Accessed February 1, 2019, https://www.opm.gov/policy-data-oversight/data-analysis-documentation/personnel-documentation/processing-personnel-actions/gppa31.pdf.

Office of Personnel Management (2015) Competitive hiring overview. Accessed February 1, 2019, https://www.opm.gov/policy-data-oversight/hiring-information/competitive-hiring/.

Office of Personnel Management (2016) General schedule overview. Accessed February 1, 2019, https://www.opm.gov/policy-data-oversight/pay-leave/pay-systems/general-schedule/.

Office of Personnel Management (2017) General employment data. Accessed February 3, 2019, https://www.opm.gov/news/releases/2017/09/opm-releases-veteran-employment-data/.

Office of Personnel Management (2018) Handbook of occupational groups and families. Accessed February 3, 2019, https://www.opm.gov/policy-data-oversight/classification-qualifications/classifying-general-schedule-positions/occupationalhandbook.pdf.

Office of Personnel Management (2019) Federal civilian employment. Accessed February 3, 2019, https://www.opm.gov/policy-data-oversight/data-analysis-documentation/federal-employment-reports/reports-publications/federal-civilian-employment/.

Partnership for Public Service (2014) Federal departures. Accessed January 14, 2019, https://www.eenews.net/assets/2014/08/14/document_pm_01.pdf.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Accessed December 3, 2019, https://www.R-project.org/.

Shoop T (2005) Do the retirement wave. Government Executive; Washington 37(21):94.

Wirth R, Hipp J F (2000) CRISP-DM: Towards a standard process model for data mining. Accessed January 15, 2019, https://pdfs.semanticscholar.org/48b9/293cfd4297f855867ca278f7069abc6a9c24.pdf.

# INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California